RESPONSE SURFACE METHODS FOR HIGH DIMENSIONAL STRUCTURAL DESIGN PROBLEMS

By

ROBERTO VITALI

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2000

A mio padre

ACKNOWLEDGMENTS

First I thank Professor Raphael Haftka, chairman of my advisory committee, for much-needed guidance during my research. He provided the funding necessary to complete my doctoral studies and constantly encouraged me to attend conferences and to publish my work in scientific journals. His help extended beyond problems encountered in the academic world. I sometimes wondered where he found the patience he always had with me. Without it this work would not have been possible.

I would also like to thank the members of my committee: Bhavani V. Sankar, Loc Vu-Quoc, Peter G. Ifju, Andre I. Khuri and Evgeni Petrov. I am grateful for their willingness to serve on my committee, their help whenever required, and their involvement with my oral qualifying examination, and their review of this dissertation.

My colleagues in the Structural and Multidisciplinary Optimization Group at the University of Florida also deserve thanks for their help, and for their many fruitful discussions. I would like to thank Melih Papila, Satchi Venkataraman, Gerhard Venter and Steen Lauridsen for being good friends also outside the walls of our office. In addition I would like to thank Leslie Ann Myers for the support that she gave me during these years.

I also would like to thank the people of the Imperial College of London, England and the University of Aalborg, Denmark for giving me the opportunity to work with them at their institutions.

Finally, I would like to acknowledge my good friend Jens Madsen, for the fun we had studying, working and playing together. If for nothing else I should be grateful to him for convincing me to stick with engineering rather than pursue an uncertain career in riveting.

TABLE OF CONTENTS

page

ACKNOWLEDGMENTS iii

LIST OF TABLES viii

LIST OF FIGURES xii

ABSTRACT xiv

INTRODUCTION 1

LITERATURE REVIEW 3
- Function Approximation Concepts 3
- Problems in Engineering Optimization based on Computer Simulations 5
- Local Approximations 7
- Mid-Range Approximations 10
- Global Approximations 13
  - Multifidelity Response Surfaces 17
  - Multiple Models in Optimization 22
  - Visualization and Response Surfaces 25
  - Parallel Computations 28
  - Numerical Noise 32
  - Outliers 37
  - Response surface based on gradient information 38

RESPONSE SURFACE METHODOLOGY 46
- Fitting an approximation to given data 46
- Evaluation of predictive capabilities 53
- Elimination of redundant parameters 56
- Design of experiments 60
  - Factorial designs 61
  - Fractional factorial designs 63
  - Central Composite Designs 65
  - Plackett-Burman designs 67
  - Box-Behnken design 67
  - Simplex design 69
  - D-Optimality 70
  - Latin Hypercube Sampling 73

Other methods for obtaining designs 77
Optimization 78

STRUCTURAL OPTIMIZATION OF A HAT STIFFENER PANEL 79
Summary 79
Introduction 80
Problem Description 83
Structural Optimization 84
Design Using PANDA2 for Initial Load Condition 86
Design Using Response Surface Techniques for Initial Load Condition 89
Buckling Response Surface Approximation 95
Stress Response Surface. 98
Optimization Using the Response Surfaces Approximations 99
Design for Updated Load Conditions 101
Approximate Buckling Analysis. 104
Minimum Weight Structure. 109
Concluding Remarks for Structural Optimization of the Hat Stiffener Panel 111

MULTIFIDELITY DESIGN OF A COMPOSITE PANEL WITH A CRACK 112
Summary 112
Introduction 112
Correction Response Surface 114
Example Problem Description 116
Finite Element Models 118
Stress Intensity Factor Calculation 119
Linear Correction Response Surfaces Approximations 120
Response surface approximation based on the low-fidelity model 123
Stress intensity constraint approximation 125
Response surface approximation based only on high-fidelity method 128
Optimization 130
Concluding Remarks 131

RESPONSE SURFACE APPROXIMATIONS INCLUDING GRADIENT INFORMATION 133
Introduction 133
Regression analysis based on function and gradient values 134
Statistical properties of gradient based response surfaces 138
Iterative procedure 140
Different forms of the estimated covariance submatrix $\hat{\mathbf{C}}$ 142
Numerical tests 143
Modeling or Bias Errors 144
Example 1. Quadratic approximation of a 4^{th} order polynomial 145
Increasing the order of the approximation 148
Example 2: Cubic approximation of 4^{th} order polynomial 149
Example 3: Ten design variable case 150
Noise Errors 152

Example 1: Approximation of 2nd order polynomial with random errors $\sigma^2 = 0.03$ 153
Example 2: Approximation of 2nd order polynomial with random errors $\sigma^2 = 0.06$ 156
Example 3: Approximation of 2nd order polynomial with random errors $\sigma^2 = 0.03$ on the function values and $\sigma^2 = 0.06$ on the gradient values 157
Example 4: Approximation of 2nd order polynomial with random errors $\sigma^2 = 0.3$ 158
Example 5: Approximation of 2nd order polynomial with correlated random errors 159
Reduction of terms in the approximations 160
Conclusion on including gradient information in constructing approximations 161

CONCLUDING REMARKS 163

LIST OF REFERENCES 165

BIOGRAPHICAL SKETCH 174

LIST OF TABLES

Table Page

1: Material properties for graphite-epoxy preforms....84

2: PANDA2 optimization problem....87

3: PANDA2 original and rounded optimum design....88

4: Design variables used in response surface approximation optimization....93

5: Candidate points for response surface construction....94

6: Structural designs used for the linear response surface....96

7: Initial optimization problem using response surfaces....100

8: Initial optimum design obtained using response surface....100

9: Structural designs near the initial optimum design that was used for the updated response surface optimization....103

10: Scale factors for combined load case buckling load factor λ_c....108

11: Minimum weight structural design....109

12: Material properties of AS4/3501 -6....117

13: Accuracy measures of the correction response surfaces β^{Lin}, δ^{Lin} response surface $K_{HF}^{0,Lin}$ obtained using only high-fidelity analyses The subscript OR and CH denote the original set of 16 points and the independent set of 48 check points, respectively....123

14: Statistical properties of quartic response surface approximation of $K_{LF}^{0,Quartic}$ obtained using 464 low-fidelity analyses. The subscript OR and CH denote the original set of 464 points and the independent set of 4032 check points, respectively....124

15: Quartic response surface approximations statistical parameters of $K_{Quartic}^{0,\beta}$ and $K_{Quartic}^{0,\delta}$ obtained from correcting K_{LF}, with β^{Lin} and δ^{Lin}. The subscript OR denotes the original set of 16 points. 127

16: Comparison of the accuracy of the stress intensity factor constraints obtained by different methods. $K_{Quartic}^{0,\beta}$ is the quartic approximation obtained from 464 β^{Lin} corrected values. $K_{Quartic}^{0,\delta}$ is the quartic approximation obtained from 464 δ^{Lin} corrected values. $K^{0,\beta}$ is the approximation obtained by multiplying the response surface β^{Lin} by the quartic approximation based on 464 low-fidelity analyses. $K^{0,\delta}$ is the approximation obtained by adding the response surface β^{Lin} to the quartic approximation based on 464 low-fidelity analyses. The subscript OR and CH denote the original set of 16 points and the independent set of 48 check points, respectively. 127

17: Response surface approximation $K_{HF}^{Quadratic}$ obtained from only high-fidelity analyses. The subscript OR and CH denote the original set of 40 points and the independent set of 37 check points, respectively 129

18: Minimum weight optimization problem for the bladed stiffened panel 130

19: Optimum designs obtained from the different stress intensity factor approximations. %Err denotes the percentage error of each approximation at its optimum (negative when conservative) 131

20: Parameters and error measures of the approximation $\hat{y}_{MinBias}$ that minimizes analytically the bias error, of the approximation $\hat{y}_g$, obtained using function and gradient values with iterative covariance algorithm, of the approximation $\hat{y}_I$ obtained using function and gradient values with an identity matrix as a weighting matrix and of the approximation $\hat{y}_f$ obtained using only function values. The underscript *f* refers to function values while *g* refers to gradient values. RMSE refers to the original nine points while *Re* refers to the analytical RMSE calculated over the entire design domain 147

21: Parameters and error measures of the cubic approximation $\hat{y}_g$, obtained from function and gradient values and the quadratic approximation $\hat{y}_f$ obtained using only function values. The underscript *f* refers to function values and *g* refers to gradient values. RMSE refers to the original nine points and *Re* to the analytical *RMSE* 150

22: Response surfaces obtained from the true function and true gradient values at 70 design points. The true function is a quartic polynomial in ten design variables. *Re* indicates the analytical *RMSE* calculated over the whole design

domain. The underscript f refers to function values and g refers to gradient values. The $Re_{g,Sum}$ is the sum of the ten Re of the gradients. Min Bias refers to quadratic and cubic polynomials that minimize the analytical bias error of the function. 152

23: Response surfaces obtained from the true function and true gradient values with added noise. The noise added had variance $\sigma^2 = 0.03$ for both function and gradient values. Errors measures Re are calculated at 101^2 check points. The underscript f refers to function values and g refers to gradient values. The number 1 or 2 refers to the design variable number. Av means average. 154

24 Response surfaces obtained from the true function values with added noise. The noise had variance $\sigma^2 = 0.03$. The errors measures Re are calculated at 101^2 check points. The underscript f refers to function values and g refers to gradient values. The number 1 or 2 refers to the design variable number. Av means average. 154

25: Response surfaces obtained from the true function and gradient values with added noise. The noise had variance $\sigma^2 = 0.06$ for both function and gradient values. The errors measures Re are calculated at 101^2 check points. The underscript f refers to function values and g refers to gradient values. The number 1 or 2 refers to the design variable number. Av means average. 156

26: Response surfaces obtained from the true function values with added noise. The noise had variance $\sigma^2 = 0.06$. The error measures Re are calculated at 101^2 check points. The underscript f refers to function values and g refers to gradient values. The number 1 or 2 refers to the design variable number. Av means average. 156

27: Response surfaces obtained using function and gradient having different noise levels. The noise added to function values had variance $\sigma^2 = 0.03$ while the noise added to the gradient values had variance of $\sigma^2 = 0.06$. The error measures Re are calculated at 101^2 check points. The underscript f refers to function values and g refers to gradient values. The number 1 or 2 refers to the design variable number. Av means average 157

28: Response surfaces obtained using only function values with added noise having variance $\sigma^2 = 0.03$. The error measures Re are calculated at 101^2 check points. The underscript f refers to function values and g refers to gradient values. The number 1 or 2 refers to the design variable number. Av means average. 158

29: Response surfaces obtained using function and gradient having the same noise levels. The noise added to function values and the gradient had variance $\sigma^2 = 0.3$. The error measures Re are calculated at 101^2 check points. The underscript f refers to function values and g refers to gradient values. The number 1 or 2 refers to the design variable number. Av means average 159

30: Response surfaces obtained using only function values with added noise having variance $\sigma^2 = 0.3$. The error measures *Re* are calculated at 101^2 check points. The underscript *f* refers to function values and *g* refers to gradient values. The number 1 or 2 refers to the design variable number. Av means average. 159

31: Accuracy of the response surfaces obtained from the true function and true gradient values with correlated added noise. Analytical *RMSE* of the approximation based only on function values, $\hat{y}_f$, of the approximation based on function values and gradients using the iterative covariance algorithm $\hat{y}_g$, and of the approximation based on function values and gradients using the identity matrix as a weighting matrix, $\hat{y}_I$ 160

LIST OF FIGURES

Figure Page

1: Visualization of the design domain as in *Knill et al.* 26

2: MODELCENTER visualization of an aircraft optimization 28

3: C_{DWave} noisy behavior from *Giunta et al.* 33

4:Variable thickness plate geometry a) and noise is FE analysis b) as in *Venter et al.* 34

5: Noisy response as obtained by *Etman et al.* for the head injury criterion (HIC) as a function of the airbag hole diameter 35

6: Finite element model a) and noisy response b) obtained by *Toropov at al.* 36

7: Full factorial design, 2^3 62

8: Factorial design, 3^3 63

9: Fractional factorial design, $3^{(3-2)}$ 65

10: Central composite design in 3 design variables 66

11: Box-Benhken Design 68

12: Simplex design 69

13: Intervals used with a latin hypercube sample (LHS) of size $n = 5$ in terms of the density function for the normal random variable x_1 74

14: Intervals used with a latin hypercube sample (LHS) of size $n = 5$ in terms of the density function for the uniform random variable x_2 74

15: Two dimensional representation of one possible latin hypercube sample (LHS) of size $n = 5$ using one normally distributed design variable x_1, and one uniformly distributed design variable x_2 76

16: Center-body region of a blended-wing-body transport airplane 80

17: Hat-stiffened skin upper cover panel of the BWB 83

18: Flow chart of the optimization history of the hat stiffened panel 85

19: Cross-section of a simple module and design variables used for the PANDA2 optimization 87

20: Division of the hat-stiffened panel into three spanwise sections 90

21: Cross-sectional geometry of a variable thickness panel 91

22: Examples of buckling modes predicted by STAGS analyses. Design point numbers refer to Table 9 105

23: Linear bifurcation buckling mode from a nonlinear prestress state of optimal design for combined loading condition 110

24: Wing panel geometry with a crack 117

25: High-fidelity finite element model of ¼ of wing panel 118

26: Quartic response surface approximation of the stress intensity factor, $K_{LF}^{0,Quartic}$ based on 464 low-fidelity analyses versus the stress intensity factor obtained from low-fidelity analysis K_{LF} 125

27: Response surfaces approximations and corresponding derivatives. Dots represent the data used a) response surface obtained from function values only $\hat{y}_f$ b) derivative of $\hat{y}_f$ c) response surface obtained from function values and gradient data $\hat{y}_g$ d) derivative of $\hat{y}_g$ 146

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

RESPONSE SURFACE METHODS FOR HIGH DIMENSIONAL STRUCTURAL DESIGN PROBLEMS

By

Roberto Vitali

December 2000

Chairman: Prof. Raphael T. Haftka
Co-Chairman: Bhavani V. Sankar
Major Department: Department of Aerospace Engineering, Mechanics, and Engineering Science

Approximation concepts are an effective approach for alleviating some of the problems associated with the direct use of modern computerized analysis techniques in an optimization environment. Response surface approximations shift the computational burden from the optimization problem to the problem of constructing the approximations. Additionally, response surface approximations filter out the noise inherent to most numerical analysis procedures, by providing smooth approximate response surface functions; and simplify the integration of analysis and optimization codes. Unfortunately response surface techniques suffer from the so-called "curse of dimensionality" and become less and less attractive as the number of design variables increases. This work investigates the use of response surface (RS) approximations in computationally expensive structural optimization problems. Two methods are proposed to extend the

applicability of response surface techniques to problems that deal with more than a few design variables.

The first approach is to alleviate computational costs by taking advantage of the simple models that were in use a generation ago when computers were much less powerful. These models, termed here as *low-fidelity* models, are less accurate. However, they can be combined with more accurate but expensive *high-fidelity* models, through the creation of correction response surfaces, to provide a good combination of high accuracy and low computational cost. Correction response surface approximations are used to solve two structural weight optimization problems efficiently: the design of the top cover panel of a blended-wing-body airplane and the design of a wing composite blade-stiffened panel subject to crack-propagation constraints.

The second approach uses response surface approximations that incorporate both function and derivative data. In computer simulations, derivatives often are available, frequently at a low computational cost. In theory the inclusion of derivatives data would allow response surface approximations to be applied to a problem with many design variables. However, derivatives are usually less accurate than function evaluations. The present study attempts to identify the trade-offs of using derivatives in constructing response surfaces through simple example.

CHAPTER 1
INTRODUCTION

Designing an engineering system is often a challenging task. Deriving a suitable mathematical formulation that accurately describes the behavior of the engineering system can be even more challenging. When this is at all possible, finding a solution to the mathematical formulation of the system might be even more difficult.

In many instances, however, it is possible to modify the original problem into a slightly different approximate one and then to solve exactly the approximate problem. The solution obtained may be viewed as an approximate solution to the original design problem. In many other instances it may be possible to solve the mathematical formulation of the original design problem using an approximation technique. The solution is once again viewed as an approximate solution to the original design problem. Of course, for a large variety of problems it is possible to solve the approximate problem in an approximate way and still obtain a satisfactory solution to the original design problem.

In this study, particular attention is paid to one approximation method called response surface (RS) technique. Response surface methodology approximates the response of a system as a function of some variables. The approximation is obtained by fitting the system response for a number of selected combinations of the control variables (design points). Response surface approximations have been shown to be accurate for a small number of parameters. However as the number of parameters increases, traditional

response surface techniques require an increasing number of evaluations to achieve acceptable accuracy. This makes them computationally intractable.

The study focuses on methods that can be applied to response surface techniques to design optimization problems with more than a few design variables. In particular the two approaches that were studied are correction response surfaces (CRS) and derivative based RS. The objective is to demonstrate the potential of these two techniques.

Chapter 2 provides a literature review on the state of the art in approximation methods applied to optimization problems. It briefly introduces the most popular approximation methods, dividing them into local, mid-range, and global approximations. Particular attention is paid to global methods in the form of RS. The section on response surface methodology, being the main focus of this work, is discussed in greater detail. The mathematical and statistical bases of response surface methodology are summarized briefly in Chapter 3. Chapter 4 applies of correction response surface to the optimization of a hat-stiffened panel made of composite materials. Chapter 5 applies correction response surface methodology to design optimization against crack propagation of a bladed stiffened panel, also made of composite materials. Chapter 6 deals with response surface based on both function and gradient data. The approach is applied to numerical problem where control on the noise and modeling errors could be enforced. Concluding remarks and proposed future research are provided in Chapter 7.

CHAPTER 2
LITERATURE REVIEW

Approximation concepts in structural optimization were introduced in the mid–seventies by *Schmit and Miura,* (1976) and by *Schmit and Farshi,* (1977). They showed that applying non linear programming methods to large structural design problems were cost effective, provided that suitable approximations concepts were introduced. Several other publications on approximation concepts soon followed as the approach attracted the interest of many other researchers.

Function Approximation Concepts

The use of approximation techniques to solve an optimization problem allows a designer to minimize the computational effort needed to find a satisfactory solution and to integrate different simulation codes if more then one code is needed during the design process. The selection of appropriate approximating functions to replace the real functions that describe an optimization problem is essential to ensure that the optimization procedure is solved correctly and efficiently.

Barthelemy and Haftka (1993) distinguished approximation methods based on the size of the domain of their validity, as local, global, and mid-range approximations.

In general, an optimization problem can be formulated as the minimization of an objective function $F(\bar{x})$

$$\begin{aligned} & F(\bar{x}) \rightarrow \min \\ & G_j(\bar{x}) \leq 0 \qquad j = 1,...,J \\ & H_k(\bar{x}) = 0 \qquad k = 1,...,K \\ & x_i^l \leq x_i \leq x_i^u \qquad i = 1,...,I \end{aligned} \tag{2.1}$$

where the vector $\bar{x} \in \Re^I$ in Eq. (2.1) defines a point in the design space Ω whose elements are the design variables. $G_j(\bar{x})$ and $H_k(\bar{x})$ are, respectively, the inequality and equality constraints that, together with the lower and upper bounds of the design variables x_i^l and x_i^u, define the feasible design domain. These lower and upper bounds of the design variables also are called side constraints.

Approximation techniques replace the problem formulated in Eq. (2.1) by a series of approximate problems. The solution of each approximate problem is called a cycle. At each cycle the original problem in Eq. (2.1) is replaced by

$$\begin{aligned} &\widetilde{F}^p(\bar{x}) \rightarrow \min \\ &\widetilde{G}_j^p(\bar{x}) \le 0 \qquad j = 1,...,J \\ &\widetilde{H}_k^p(\bar{x}) = 0 \qquad k = 1,...,K \\ &x_i^{l,p} \le x_i \le x_i^{u,p} \qquad i = 1,...,I \end{aligned} \tag{2.2}$$

where p is the cycle number, and $\widetilde{F}^p(\bar{x})$, $\widetilde{G}_j^p(\bar{x})$, and $\widetilde{H}_k^p(\bar{x})$ are approximate expressions for the objective function, inequality constraints, and equality constraints, respectively, and the side constraints satisfy the condition

$$x_i^{l,p} \ge x_i^l \quad \text{and} \quad x_i^{u,p} \le x_i^u \tag{2.3}$$

The lower and upper bounds of the design variables change from cycle to cycle in order to limit the search for the solution to a region where the approximations are valid. The number of cycles needed to reach an acceptable design may vary from one, in the case of very accurate global approximations; to thousands, in the case of local approximations that are valid only in small regions of the design domain.

Problems in Engineering Optimization based on Computer Simulations

Numerical simulations enable engineers to perform detailed analyses of complex systems. However, the use of such simulation tools in design optimization is associated with several problems. A numerical simulation provides a designer with only a single point in the design space rather than an overall view of the system behavior within the design space. This is in contrast with the pre-computer era, when algebraic expressions and design charts furnished a global view of the system behavior in the design space. The algebraic expressions and design charts allowed the designer to spot promising design regions and structural configurations. However analysis methods available in the pre-computer era were not accurate enough for the design of complex structures. It would be beneficial if it were possible to generate a global view of the design space by producing algebraic expressions based on accurate numerical simulations.

A second problem with computer simulations is that human errors and numerical noise usually are present when dealing with large, complex, and cumbersome simulation software such as finite element codes. Numerical noise can be caused by mesh generator procedures that update the element mesh as the design changes in order to maintain required accuracy. Other sources of noise include less than fully converged iterative procedures and round-off errors. Because the number of elements and grid points in a finite element (FE) model is discrete by nature, finite jumps in the evaluated response can be introduced when an infinitesimal change in the FE model leads to adaptive meshing that changes the number of nodes and elements or the topology of the mesh. Human and numerical errors of considerable magnitude are easily located and corrected, so normally the remaining errors have little influence on the accuracy of the results. However, the remaining numerical noise may cause non-smooth behavior of the response, trapping the

optimizer in spurious local optima and making it difficult to find the best design (*Giunta et al.*, 1994).

When analyzing a complex structure using a finite element code, a source of noise is introduced implicitly in the discretization procedure. *Van Keulen et al.* (1997) performed an extensive study of the influence of numerical noise on the optimal solution of several structural design problems for different discretization densities. They showed that high levels of noise can be tolerated at the early stages of optimization, but as the optimization progresses and more information about the behavior of the response is gathered, the level of noise that can be tolerated is reduced.

Optimization is a procedure that searches through the design domain for the optimum design. During this search, the optimization algorithm continuously calls the analysis code and, based on the data obtained, it decides on the next design point to evaluate. This requires an interface between the analysis and optimization programs. Some general-purpose finite element programs, like NASTRAN, come equipped with an optimization algorithm and an interface between optimization and analysis. However, often the optimization code and the analysis code are separate, or one might want to use an optimization algorithm different from the one furnished with the analysis code. In these situations the user must write an interface between the two codes, possibly a cumbersome and frustrating task.

Finally, some optimization problems such as global optimizations or stochastic optimization require a large number of analyses and, except for simple problems; such a large number of analyses are computationally impractical.

As we show, the use of approximations, especially global approximations, helps alleviate these four problems. Approximations can provide the designer with a global

view of the design space, filter out numerical noise, help in interfacing analysis and optimization software, and permit us to perform global or stochastic optimization.

Local Approximations

Local approximations are probably the most commonly applied approximations in structural optimization. Generally, they are based on the function value and sensitivities at one point of the design domain. Much literature has been written over the years on these types of approximations (e.g., *Barthelemy and Haftka*, 1993). As their name suggests, local approximations are only accurate in the vicinity of the point at which they are generated.

A distinction should be made between local function approximation and local problem approximation. Local function approximations are variations on Taylor series expansions, while local problem approximations try to reduce the size of the active constraint set or the design variable set.

One of the first robust optimization algorithms was the simplex algorithm (*Dantzig,* 1963) for the solution of linear optimization (known as linear programming) problems. It was only natural that the first approximation methods developed used linear programming. Nonlinear optimization problems were linearized about a point at each optimizaton cycle and the simplex method was used to find the optimum in the small neighborhood. The cycle was repeated until convergence leading to the so-called sequential linear programming (SLP) method.

For a linear approximation, the functions $\widetilde{F}^p(\bar{x})$, $\widetilde{G}_j^p(\bar{x})$, and $\widetilde{H}_k^p(\bar{x})$ in Eq. (2.2) are linear Taylor expansions. For example

$$F(\bar{x}) \approx \widetilde{F}^p(\bar{x}) = F(\bar{x}_0) + \sum_{i=1}^{I}(x_i - x_{0i})\left(\frac{\partial F(\bar{x})}{\partial x_i}\right)_{\bar{x}_0}$$

where $\bar{x}_0$ is the design point about which linearization is performed.

It was realized soon that for some problems more accurate approximations could be obtained for the same computational cost. The first to come was probably the reciprocal approximation, where the linear approximation of $F(\bar{x})$ is replaced by a linear approximation in $F(\bar{y})$ where

$$y_i = \frac{1}{x_i} \tag{2.4}$$

This substitution usually helps because in many structural design problems the design variables are truss cross-sectional areas, moment of inertia of beams and thicknesses of plates. Displacements and stresses of statically determinate structures are linear in the reciprocals of these design variables, and for statically indeterminate structures approximate linearity often holds. One problem with the reciprocal approximation is that it becomes unbounded when the design variables approach zero. One way to avoid the problem was to make use of the modified reciprocal approximations (*Haftka and Shore*, 1979). Another first-order approximation method, called the conservative approximation, was introduced by (*Starnes and Haftka,* 1979). This method is a hybrid form of the linear and reciprocal approximations methods that is more conservative than either. The original or reciprocal approximations are used according to the sign of the product.

$$x_{oi}\left(\frac{\partial F(x)}{\partial x_i}\right)_{x_o} \tag{2.5}$$

One advantage of the conservative approximation is that it renders the problem convex and therefore is guaranteed to have only one global minimum. Several researchers proposed different approximation methods that had the desirable property of

convexity; the most successful one may be the method of moving of asymptotes introduced by *Svanberg* (1987). However the reciprocal and the conservative methods as well as other convex approximation destroy the linearity of the problem, thereby rendering impossible the use of SLP.

Higher derivatives can be used in local approximations, but the increase in accuracy often is not worth the computational cost of the additional derivatives. Quadratic approximations primarily have been used in solving eigenvalue problems. However, *Murthy and Haftka* (1988) showed that for almost the same computational cost it is possible to obtain a third order approximation by using a linear approximation to the eigenvectors in the Rayleight quotient. Often the high cost of obtaining second derivatives dictates a compromise leading to the use of only the diagonal elements of the second derivatives matrix (*Fleury,* 1989).

Using approximations, it is not necessary to approximate all response quantities considered in the optimization process. Using the process of local problem approximation we can ignore constraints that are not critical or near critical for the current cycle in the optimization process. This approach, referred to as constraint deletion or constraint screening, was first introduced by *Haug and Aurora* (1979). It should be noted that constraint deletion is used only to reduce the computational effort but that it does not affect the final optimum achieved. For a more complete overview of the process see *Vanderplaats* (1999). Moreover when dealing with complex FE models, many elements in a small region of the structure have approximately the same stress. It is possible to use the most critical element to represent the response in that region. This approach, known as regionalization, was introduced by *Schmit and Miura* (1976). The result of these two simple techniques is that while the structure may have several thousand constraints, the

sensitivity of only a small fraction of them has to be evaluated, leading to considerable savings in computational cost.

Local approximations have been successfully applied to numerous problems but they have some drawbacks. First, they do not provide the designer with a global view of the design space. Second, local approximations provide only limited protection against numerical noise. Finally, local approximations are not well suited for parallel computing without non-trivial software changes.

Mid-Range Approximations

Mid-range approximations are valid in a region of the design space that is considerably larger than local approximations but smaller than the entire design space. Unlike local approximations, data from more than one design point are used to construct mid-range approximations. In local methods, a new approximation is constructed at a new design point, discarding previous analyses. Some mid-range approximation methods work in a similar way, but now at each cycle more than one point is used to obtain an approximation valid in a region of the design domain. Mid-range methods use data from previously generated design points to enhance the accuracy of the approximations. In the literature two types of mid-range approximation methods that use the previously generated data can be distinguished. The first type uses data from individual points along the optimization path in the design space. Such approximations are called single-point path (SSP) approximations. Approximations derived from data computed in clusters of design points along the optimization path are called multiple point path (MPP) approximations.

In SPP approximation, the design points used for the approximations are the solutions of the previous optimization cycles. *Vanderplaats* (1979) developed Taylor

series expansions about the current design point in optimizing airfoils using function values from all previous design points. Two and three-point approximations were proposed by *Haftka et al.* (1987). The authors generated approximations based on the projection of a design point on the line connecting two points, or onto the plane connecting three points, approximating the functions by Hermite polynomials at the projection point. While tests showed improvement over local linear approximations, reciprocal methods still gave better results. In addition the results indicated that the approximations were good for interpolations, while the improvements in accuracy were marginal for extrapolations. *Fadel et al.* (1990) presented a two-point exponential approximation in terms of the intervening variables $y_i = x_i^{p_i}$, $i = 1,\ldots,\mathrm{I}$. The exponent p_i for each design variable was determined by matching the derivatives of the approximate function with the previous data point gradients.

Grandhi and co-workers (e.g., *Wang and Grandhi*, 1994, *Wang and Grandhi,* 1995 and *Xu and Grandhi,* 1999) constructed a family of algorithms based on intervening variables called TANA, where the intervening variables were defined as in *Fadel el al.* as $y_i = x_i^{p_i}$ for $i = 1,\ldots,\mathrm{I}$. For example, the TANA3 algorithm of *Xu and Grandhi* (1999) uses a second-order Taylor series expansion in terms of the intervening variables, with a diagonal Hessian matrix. The exponent of each design variable and the diagonal second-order terms are evaluated by matching the derivatives and the function value of the approximate function with the previous function and gradient data. In other words, unlike the approach followed by *Fadel et al.,* TANA3 matches the function value and the gradient of a previous design point rather than only the gradient. *Rodriguez et al.* (1998) implemented an algorithm that is based on an approximation of the Hessian matrix. The authors approximated the Hessian matrix for their problem by matching the function and

gradient values of the response surface approximation with the values obtained from numerical experiments in a region of the design space at additional design points around the current one.

An improvement to local methods, known as the accumulated function approximation, was developed by *Rasmussen* (1996). Starting from an arbitrary design point, a local approximation is constructed and a possible optimum point is found in a design space larger than the space of validity of the local approximation. Based on the function value and possibly the gradients at the starting point and newly obtained point, a new approximation is found for the design space. The process allows including information from more points as they become available. A weighting procedure is used so that points closer to the last point have more weight in defining the approximation. *Rasmussen* (1996) showed that even a considerable level of noise in the sensitivities can be tolerated by this method. No noise in the function values was considered.

The MPP methods generate one or more extra design points around each solution of an optimization cycle, enabling the approximations to use these additional points. The additional points can be chosen according to design of experiments discussed in *Khuri and Cornell* (1996). *Free et al.* (1987) used a central composite design (CCD) scheme to select design points used to construct low-order polynomials in a region of the design domain. They reported competitive efficiency with conventional optimization algorithms when no noise is present. When noise is present, the method was more efficient than conventional algorithms. A general MMP concept was formulated by *Toropov et al.* (1999) in which any function $\widetilde{F}(\bar{x})$ in Eq. (2.2) can be used as an approximation function. These models can be fitted to a function $F(\bar{x})$ by means of weighted least-squares method as

$$\min \sum_{i=1}^{I} \left(w_i \left(F(x_i) - \tilde{F}(x) \right) \right)^2 + \sum_{j=1}^{n} w_{j,i} \left(\frac{\partial F}{\partial x_i} - \frac{\partial \tilde{F}}{\partial x_i} \right)^2_{x_i} \qquad (2.7)$$

where w_i contains the weight factors expressing the relative contribution of each design point. Note that in this approach the use of first-order derivatives is not mandatory. The procedure to construct function approximation is comparable to global methods. The difference is that the region where the approximation is used, bounded by move limits, moves through design space as the optimization progresses. *Van Keulen et al.* (1995) used the multi-point-path-approximation mid-range methods successfully to solve of a shape optimization problem. A similar approximation concept was applied successfully by *Etman et al.* (1996).

Global Approximations

For some problems, it could be profitable to construct a global approximation instead of a local or mid-range one. One of the most commonly used methods for global approximations is the response surface (RS) technique, originally developed for the planning and analysis of physical experiments (e.g., *Box and Draper*, 1987). Because of similarities between physical experiments and computer simulations, response surface techniques have been used to build explicit approximations to structural response from numerical experiments in the form of finite element simulations.

Response Surface Techniques

When dealing with complex problems the true response or performance measures as a function of the design variables usually is not known explicitly. In these situations design optimization based on experiments may be the only tool a designer has.

For example, a farmer may want to maximize the yield per acre of crops by varying the quantity of two fertilizers. The true value of the yield as a function of the two fertilizers is not known, and it may not be derived explicitly . The farmer in this case can perform several experiments on different plots using different quantities of the two fertilizers and can record the yield for each combination of the fertilizers. Once the data from the experiments has been gathered, an analytical expression relating the two design variables (fertilizers) and the response (yield per acre) can be fitted to the data to generate a response surface as a model of yield. Most commonly, linear or quadratic polynomials are used. However, noise in the data may be introduced due to different skills of workers tending the plots or inaccuracies in measuring fertilizer amounts or yields. The farmer would probably like to conduct a large number of experiments to reduce the effect of the noise, but at the same time he or she would like to keep the number of experiments small to minimize cost.

Response surface techniques help the farmer to select combinations of the two fertilizers to try in order to gain the maximum information from the minimum number of experiments. In addition, after the analytical expression of the response is fitted (a process called regression) response surface methodology helps the farmer establish whether the relationship found is statistically meaningful. Can the farmer trust the relationship found to predict the yield of crops for all the possible combinations of the fertilizers? What is the expected error in the yield prediction? Can the prediction of the yield response be enhanced? If more experiments were needed, which combinations of the two design variables would be desirable? Are some of the experiments suspect (possibly due to human error) and should be removed as "outliers"? Response surface techniques answer these questions and deal with these issues.

Response surface methodology initially was developed and described in a paper by *Box and Wilson* (1951), and further investigated in subsequent papers *Box and Hunter* (1957) and *Box and Draper* (1959). It can be defined as " a group of statistical techniques for empirical model building and model exploitation" (*Box and Draper,* 1987), or similarly as "a collection of statistical and mathematical techniques useful for developing, improving, and optimizing processes" (*Myers and Montgomery,* 1995).

Papers and the books devoted to the topic of RS, *Myers and Montgomery* (1995), *Box and Draper* (1987) and *Khuri and Cornell* (1996) focus mainly on physical experiments and how to model the results from field tests in order to achieve maximum information from the data collected. *Hill and Hunter* (1966) present several examples of early applications of response surface methodology in a review paper.

Because of the increasing commonality between physical and numerical experiments, these techniques or methodologies have been extended successfully to engineering optimization based on computer simulations. First, like physical experiments numerical simulations such as finite element analyses, can be expensive in terms of time and cost, especially in terms of model preparation and validation. Second, in performing a large number of experiments needed for optimization there is usually an advantage in performing them in batches. This is true for the farmer who may need to wait for several months for the results of an experiment. It is true for the engineer in a laboratory because necessary equipment may not always not be available, and therefore experiments need to be scheduled in batches when the equipment is available. Similarly, multiple numerical simulations can be performed at off peak hours, at night or during weekends when computers are rarely used. In addition, the availability of parallel computers naturally allows analyses to be performed in batches on different nodes of a parallel computer

rather than one at the time (*Balabanov et al.,* 1996). Third, like physical experiments, complex numerical simulations often acquire numerical noise from different sources. However some differences remain between physical and computational experiments, perhaps the most important one is that the repetition of a numerical experiment usually gives the same result, while this is far from true for a physical experiment.

Response surface methodology can be used to construct global and mid-range approximations to functions in structural optimization. Traditionally, linear and quadratic polynomials are used. Because structural optimization requires expensive function evaluations, it is important to construct accurate function approximations so that rapid convergence may be achieved. *Toropov* (1989) investigated the role of the choice of approximating functions on the accuracy of the approximations. Different techniques to find the region of interest containing the optimal design, and techniques for finding more accurate approximations also were presented by *Roux et al.* (1998). *Roux* and co-workers solved the weight minimization problem for two-, three-, and ten-member trusses using response surface methodology and compared their results with solutions in the literature (*Arora,* 1989).

The authors found that the accuracy of the response surface is influenced strongly by the number and the location of data points used to construct the response surface approximation. However, ultimately the accuracy that can be achieved is limited by the choice of the approximating functions and the design domain size. For a successful use of response surface methodologies, the authors advocate as the most important features the use of engineering knowledge in choosing the approximating functions and the use of as many constraints as possible in order to reduce the size of the design domain. This so-called reasonable design space approach is discussed further in the next section.

Venter et al. (1996) demonstrated the use of dimensional analysis for reducing the dimension of the design domain in their weight-optimization of a stepped plate. The authors reduced the number of design variables in their response surface approximation by using variables that are intrinsic to the problem, which were identified by nondimensional analysis of the governing equations. In contrast with *Roux et al.*, and *Venter et al.* used higher-order polynomials (cubic and quartic instead of the more traditional quadratic) to achieve greater accuracy for the stress and buckling constraints.

Several other examples of higher-order polynomials are available in the literature. *Van Houten et al.* (1995) for example used a fifth order polynomial response surface to approximate the drag coefficient in an aerodynamic shape optimization as a function of two design variables. Unfortunately response surface techniques that use higher-order or even quadratic polynomials become difficult to implement as the number of design variables increases since the number of data required to fit the response surface becomes prohibitively large. For problems with four or more design variables *Kok et al.* (1998) advocate the use a series of local linear response surface approximations until convergence to an optimum is reached. However this approach suffers from many of the disadvantages of the local approximations methods. In order to be able to use response surface techniques for high-dimensional problems another use of response surface called "multifidelity response surface methodology" is gaining popularity.

Multifidelity Response Surfaces

Most applications of response surface methodologies use a region in a design space that is a box defined in terms of lower and upper bounds of the design variables. However, unless it is possible to analyze designs at least at all the vertices of the box, the

process of approximation inside the box can entail extrapolation rather than interpolation, with an attendant of loss of accuracy. With an n-dimensional box having 2^n vertices, it becomes impractical to evaluate the designs at all the vertices for values of n of the order of 10 or more. Moreover, response surface approximation may not guarantee sufficient accuracy even when limited to interpolation since complex response patterns are often difficult to approximate as low order polynomials. A different way of using response surfaces is gaining popularity as a solution to this problem. Inexpensive lower-fidelity models often are available to the designer, and strategies can be devised that use the lower fidelity methods to produce accurate approximations at reasonable computational cost . The combined use of lower and higher-fidelity methods and response surfaces is referred to as "multifidelity" methods.

Lower-fidelity models can be obtained either by simplifying the analysis model (e.g., by using a coarser FE mesh discretization) or by using a simpler model (e.g., finite strip instead of FE). Usually the lower-fidelity models vary with the design variables in a manner similar to the high-fidelity model, thus providing a good basis for the high quality approximations. Some of the multifidelity methods use the high-fidelity models to "correct" the lower-fidelity model rather than to approximate directly the expensive higher-fidelity model. *Toropov and Markine* (1996) classify multifidelity response surface methods as multiplicative or additive. For multiplicative cases, correction response surfaces are used to multiply the low-fidelity result. In the additive cases the correction response surface is added to the results obtained from the low-fidelity method.

In this spirit, *Mason et al.* (1994) used quadratic correction response surface techniques to design channel frames subject to stress constraints where the lower-fidelity models used 2-D FE models, and high-fidelity models used 3-D FE models. A low-

fidelity model also can be a closed-form approximation. *Vitali et al.* (1997) used closed-form beam and plate solutions to predict stresses and buckling loads in a hat-stiffened panel, using the STAGS (*Rankin et al.,* 1998) FE code for correction. Constant response surface were constructed for the ratios of the stresses or buckling loads from high-fidelity and low-fidelity models at pre-selected design points. Stresses and buckling loads for other points were approximated by multiplying the lower fidelity results by the response surface.

Instead of using a constant response surface, *Venkataraman and Haftka* (1998) used a linear correction response to predict the buckling loads of cylindrical panels. The authors used the PANDA2 (*Bushnell,* 1987) program as their low-fidelity model and the STAGS FE program as their high-fidelity model. PANDA2 is a panel design program that uses closed form or Rayleight-Ritz type solutions to obtain failure loads. A quadratic response surface in five design variables was created by using many PANDA2 analyses and then a linear correction response surface approximation was fit to the ratios of the STAGS analyses and the PANDA2 analyses at few selected design points. Buckling loads for designs in the design domain then could be obtained by multiplying the quadratic response surface approximation and the linear correction response surface. The buckling load approximations were used as constraints for weight optimization of cylindrical shells.

Balabanov et al. (1998) compared the results obtained using both the multiplicative and the additive correction response surface approximations in predicting optimum wing bending material weight of a High Speed Civil Transport (HSCT) as a function of the aircraft geometry. The approximations used thousands of structural optimizations of computationally inexpensive coarse finite element models to produce a

quadratic response surface model in the 29 design variables, and then about a hundred structural optimizations of refined FE models were used to construct linear correction response surfaces approximations. They found that the linear correction response surfaces reduced the error of the quadratic response surface approximations based on coarse models by more than half. Moreover, most of the correction could be obtained with only a constant correction with linear terms adding only marginal improvements. The additive correction response surface performed better than the multiplicative one. However, the authors suggested that the advantage of the additive correction response surface is not a general rule but possibly is a feature of the specific problem that they considered.

A different strategy in combining low-fidelity and high-fidelity models was followed by *Kaufman et al.* (1996) and *Balabanov et al.* (1999) when searching for the optimal takeoff gross weight for the HCST. This strategy is called "reasonable design space approach". The reasonable design space approach seeks inexpensive constraints that eliminate from consideration large portions of the design box defined by the upper and lower limits on the design variables. Such constraints may come from simple geometric constraints that prevent combinations of design variables resulting in unreasonable geometry configurations for the aircraft. Low-fidelity methods were used to estimate the performance of candidate designs via aerodynamic constraints and performance constraints. The low-fidelity methods may not have provided accurate estimates of performance but they could identify designs with such inferior performance that they could be discarded even based on very approximate results. The points that corresponded to unreasonable aircraft configurations were not simply removed but they were moved toward the center of the design space until geometric, performance and aerodynamic constraints were met based on the low-fidelity method. The ratio of the

volume of the initial design space box to the volume of the final reasonable design space for the 25 design variables considered was about 10^{14}. After the reasonable design space was obtained moving the unreasonable design points tower the center of the design space using the lower fidelity analysis about 1000 high-fidelity analysis were performed and a response surface created using only the high-fidelity data. Using this last response surface the authors were able to find an aircraft configuration that was 18.3% lighter than the baseline configuration.

Low-fidelity methods can also be used to reduce the number of terms in a response surface and then the reduced response surface can be fitted using a smaller number of high-fidelity analyses. This approach was followed by *Knill et al.* (1999) to efficiently implement aerodynamic prediction from Euler solutions into a multidisciplinary design optimization study of the HSCT. Statistical techniques were applied to linear theory response surface models to reveal the terms that had a significant impact on calculating the desired aerodynamic quantities. Terms that had little or no effect could be removed without degrading the accuracy of the response surface models. Only important terms were retained when the reduced-terms response surface models were constructed using Euler solutions with viscous drag correction, thereby requiring fewer computer fluid dynamics (CFD) analyses to compute the coefficients of the models. The reduced-term models were created as multiplicative corrections of the full-term linear theory response surface models. In such approach there was the possibility that some important non-linear effect could be missed but this did not seem to happen. By using the low-fidelity methods to discard non important terms the authors could save 255 CPU hours out of 392 required for the full-term response surface while maintaining the same level of accuracy.

Multiple Models in Optimization

The increasing complexity of engineering systems has sparked increasing interest in multidisciplinary optimization (MDO). MDO can be described as a methodology for the design of systems where the interaction between several disciplines must be considered, and where the designer is free to significantly affect the system performance in more than one discipline. Two are the main challenges of MDO: computational expense and organizational complexity. These challenges are summarized in the following based on *Sobieszczanski-Sobieski and Haftka*, (1997) paper.

The interdisciplinary coupling inherent in MDO tends to present additional challenges beyond those encountered in a single-discipline optimization. It increases computational burden, and it also increases complexity and creates organizational challenges for implementing the necessary coupling in software systems. Since solution times for most analysis and optimization algorithms increase at a superlinear rate, the computational cost of MDO is usually much higher than the sum of the costs the single-discipline optimizations for the disciplines represented in the MDO. Additionally, even if each discipline uses linear analysis methods, the combined system might require costly non-linear analysis. For example, linear aerodynamics may be used to predict pressure distributions on a wing, and linear structural analysis may be then used to predict displacements. However the dependence of the pressure on the displacements may not be linear. Finally, for each disciplinary optimization we may be able to use a single-objective function, but for the MDO problem we may need to have multiple objectives with an attendant increase in optimization cost.

MDO also faces formidable organizational challenges. The analysis codes for each discipline have to be made to interact with one an other for the purpose of system

analysis and system optimization. Data transfers among modules of an MDO problem correspond to the internal couplings of the system. These data transfers may require data processing that may become a costly overhead. For example, if the system is a flexible wing, the aerodynamic pressure reduced to concentrated forces at the aerodynamic model grid points on the wing surface has to be converted to the corresponding concentrated loads acting on the structure finite-element nodal points. Conversely, the finite element nodal structural displacements have to be entered in the aerodynamic model grid as shape corrections. The volume of data transferred in such couplings affects efficiency directly in terms of I/O cost. Additionally many solution procedures require the derivatives of these data with respect to design variables, so that a large volume of data also increases computational cost. To decrease these costs response surface approximations can be used.

In the above wing example one may represent the pressure distribution and the displacement fields by response surface approximations, where the approximating functions constituting the response surface are chosen by the designer beforehand based on engineering knowledge. In this way only the coefficients of the approximating functions have to be transferred instead of the large volumes of the discrete load and displacement data. An example of such a condensation for supersonic transport design was reported in *Unger et al.* (1992).

Direct coupling of the optimizer to a multidisciplinary analysis may be impractical for various reasons. First, for moderate to large number of design variables, the number of evaluations of objective function and constraints required by the optimizer is high. Often we cannot afford to execute such a large number of exact MDO analyses in order to provide the evaluation of the objective function and constraints. Second, often the different disciplinary analyses are executed on different machines, possibly at

different sites, and communication with a central optimizer can become impossible. Third, some disciplines may produce jagged or noisy responses as a function of the design variables (*e.g., Giunta et al.* (1994). For all of these reasons often the optimizer is connected to easy-to-calculate approximations of the objective function and constraints. Response surface approximations are particularly suitable for this task. For MDO problems response surface techniques also provide a convenient representation of the data form one discipline to other disciplines and to the system. *Sellar et al.* (1996) designed an autonomous hovercraft where four disciplines used four different codes that were made to communicate through response surface approximations.

Response surface approximations are not only used as a mean for communication between different disciplinary codes, but can serve also as a communication tool between codes of the same discipline representing multiple models in that discipline. For example, simple but flexible interface between the global and the local design optimization codes is often needed when designing a complex structure, and response surface can serve for this purpose. *Ragon et al.* (1997) approximated the optimal panel weight obtained from the local design code with an response surface for designing wing structure. This approximation was then used as a constraint in the global wing approximation. *Liu and Haftka* (1998) used a similar approach for transmitting optimal buckling loads from the panel level to the wing level. *Venkataraman and Haftka* (1997) approached a similar problem from the opposite end when designing the structure of a laminated composite liquid hydrogen tank. Approximations in the form of response surface were constructed for the global buckling and the bowing displacement from the global analysis and then the response surface approximations were used as constraints in local optimization.

Response surface approximations have recently gained popularity as a simple way to facilitate communications between specialists on the design team. In this sense, these techniques are becoming one of the means to meet the organizational challenge of MDO (*Tai et al.,* 1995 and *Giunta et al.*, 1995). The use of response surface approximations in general often allows dealing in a better way with organizational boundaries. The approximations used for each discipline can be generated by specialists of the discipline. The specialist can tailor the approximation to special features of that discipline and to the particulars of the application. When response surface are used, the creation of the various disciplinary approximations can be performed ahead of time, minimizing the interaction of the optimization procedure with the various disciplinary software.

Visualization and Response Surfaces

Global approximation methods in general and response surface techniques in particular offer the designer a global picture of the problem, pointing out tradeoffs and particularly promising regions within the design space. Visualization techniques help the designer take advantage this information. There is a general agreement that visualization has an important role in computational science and engineering, allowing scientists and designers to gain an understanding that is otherwise not possible. Coupling visualization techniques and response surface can help in acquiring a deeper understanding of the problem at hand. *Knill et al.* (1999) used response surface techniques to optimize the takeoff gross weight (TOGW) of an HSCT aircraft. Different aircraft configurations were obtained using 10 design variables. The authors found that depending on the starting points their optimization converged to two slightly different HSCT configurations that were *2000 lb.* apart. The question of why the optimizer was not able to always reach the lowest TOGW configuration was sought by visualizing the design space. Figure 1 shows

a section of a two-dimensional slice through the multidimensional design space using the first and the second optimum found by *Knill* and co-workers, and a sub-optimal feasible point.

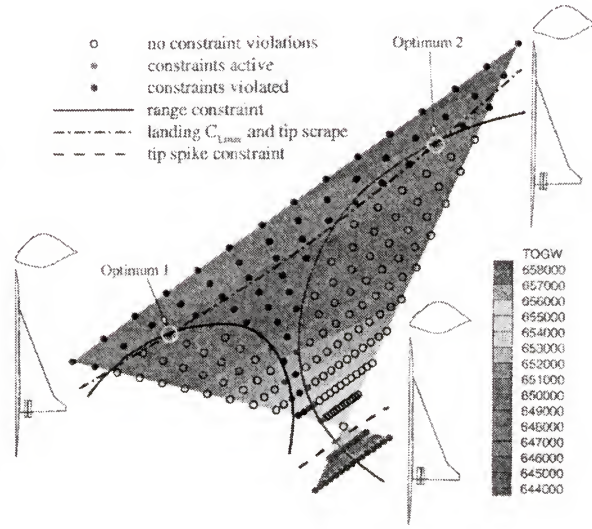


Figure 1: Visualization of the design domain as in *Knill et al.*

The remaining points in the plane were created by linear combinations of the three points. In Figure 1 the circles represent design points. Open circles represent feasible points while filled circles represent points at which the constraints are violated. The value of the objective function obtained from a response surface in the 10 design variables is indicated by the shaded contour plots. The curved lines represent the boundaries of four constraints. Prior to creating the visualization, it was not recognized by *Knill et al.* that the constraints break the design space into disjoint (at least in some hyper-planes) regions of feasible points. Only after the visualization of the design domain and objective function in a sub-domain around the two optima was obtained, it was possible to understand why the optimum depended on the starting point. Visualization in conjunction with response surface and optimization provide understanding of the

optimization process and the tradeoffs involved. They also have the potential to provide guidance by an experienced engineer when the optimizer experiences difficulties.

The HSCT aircraft is also the test bed for a specific visualization software, VIZCRAFT (*Goel et al.,* 1999) that is being developed at Virginia Tech at the MAD Center. VIZCRAFT is based on two basic tools to better manage the information available. The first tool permits the user to quickly evaluate the quality of a design with respect to its objective function and constraints. Each time the value of a design variable is modified, the HSCT planform is immediately updated and displayed to reflect the new geometry, and so are the value of the objective function and of the constraints. However, in order to have a quick prediction of the objective function and of the constraints, the VIZCRAFT program does not perform any additional analysis but relies on response surface based on previously calculated data. A costly evaluation of the objective function and constraints by CFD models of the aircraft is performed only if the user explicitly requests it. When the new evaluation requested by the user is performed then the results are added to the database on which the response surface for quick calculations are based on. Parallel coordinates is the second tool used by VIZCRAFT to visualize the results of several HSCT configurations in a high dimensional space. This technique allows the user to visualize several different HSCT design variable, constraints and objective function vectors at the same time and enables the user to recognize potential promising designs.

Malone and Woyak (1999) presented a visualization software, called MODELCENTER (*Malone and Woyak,* 1998), that help aircraft designers to combine analyses from several design disciplines into one model, and visually observe their interaction. As the optimization proceeds the designer can halt the optimization and display the aircraft geometry on the screen with a simple click. Moreover charts in two

and three dimension of any parameter combinations can be displayed so that the designer can follow the optimization as it progresses and perhaps visually detect if some anomalies occur. In Figure 2 a snap shot of the visualization as it is furnished by MODELCENTER is presented.

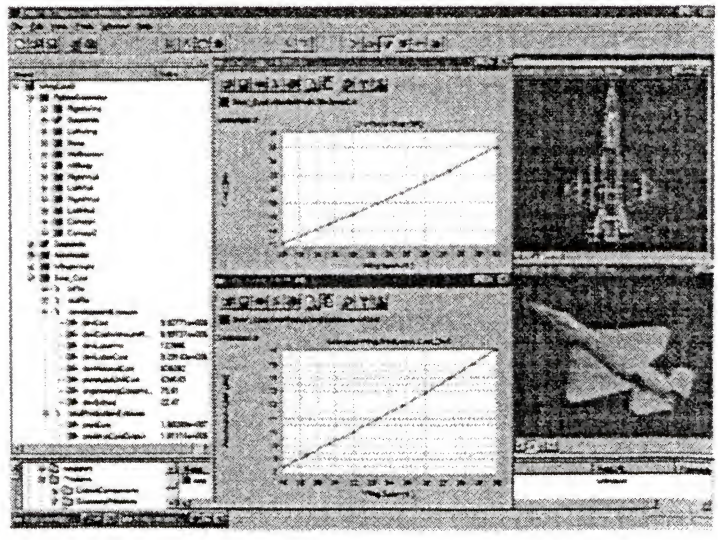


Figure 2: MODELCENTER visualization of an aircraft optimization

The potential of the software developed was shown in the cost study of a fighter. Very simple aerodynamic analysis tool were used to estimate the lift and drag performance of the wings and a response surface was used to estimate the development and manufacturing cost of the aircraft. The software used the knowledge about the components in a "smart" way, updating only the information about the components that were likely to have a grater impact on the cost of the aircraft.

Parallel Computations

Modern aerospace vehicle design requires the interaction of multiple disciplines, traditionally processed in a sequential order. MDO is evolving towards methods capable of replacing the traditional sequential methodology by concurrent algorithms, with both

an overall gain in product performance and a decrease in design time. Two methods are available to the designer when trying to exploit parallelism in optimization, coarse-grain and fine-grain parallelism. Coarse-grain parallelism involves the concurrent execution of multiple independent function evaluations while fine-grain parallelism involves computing the basic computational steps of an algorithm (i.e. the internal algebra) in parallel. In general coarse-grain parallelism requires very little inter-processor communication while the fine-grain parallelism involves much more data transfer.

An MDO paradigm based on coarse-grain parallel computing, variable-complexity modeling and multipoint response surface approximations was presented by *Burgee et al.* (1996). This paradigm interleaves the disciplines at one level of complexity and processes them hierarchically at another level of complexity, achieving parallelism within the disciplines rather then across disciplines. The coarse-grained parallelism was chosen over the fine-grained parallelism because it allows orchestrating the process with minimal knowledge of the individual disciplines involved: aerodynamics and structures. The overall parallelization task had two distinctive sub-tasks: parallelizing the individual steps of the MDO problem and linking everything together so that the entire MDO becomes a single parallel problem. The last sub-task is particularly time consuming because an MDO problem typically involves programs in several languages, extensive I/O for some program components, manual transfer of data between formats and between machines, and human intervention to interpret the output and tune the input parameters.

Burgee and co-workers used parallelization extensively while trying to optimize the TOGW of an HSCT aircraft. First a master-slave paradigm was implemented such that each node of the parallel computer performed its own optimization to find a D-Optimal set of points where to construct aerodynamic and structural RS. Then the coarse-

grained parallelization of the aerodynamic analysis modules made use of a master-slave paradigm on the Intel Paragon, whereby one designated master node controlled the data transfer and the I/O of the remaining nodes. Once all the analyses were obtained a response surface approximation for the aerodynamic response was constructed as a function of the design variables.

The parallelization of the structural analyses performed by *Burgee et al.,* followed a similar master-slave paradigm, where the structural code used was GENESIS (*VMA Engineering,* 1997). Structural analyses were performed in parallel on the Intel-Paragon at some D-Optimal points previously chosen and then a response surface approximation for the structural response was constructed. However, at first the speed-up achieved could not be increased over 2.3 no matter how many nodes were used. The speed-up is defined as the ratio of the computation time obtained with a parallel computer to the time required to solve the problem the same problem with only one processor. In theory the maximum speed-up achievable should be equal to the number of nodes used in the parallel computation. The reason of this poor result even when 20 nodes were used, was due to the huge amount of I/O that GENESIS required. For this reason, GENESIS was modified in order to diminish the volume of I/O operation and ultimately a speed-up factor of 11.7 when using 20 nodes was achieved. This was a substantial improvement but it was still not totally satisfactory. In order to achieve greater speed-ups the authors recommend further modification to the GENESIS program in order to reduce even more of I/O operations needed. Similar results were reported by *Kaufman et al.* (1996) in their study of the TOGW of an HSCT aircraft using parallel computers.

Krasteva et al. (1999) continued the studies of *Burgee et al.* using different schemes for the parallelization process in order to achieve better speedups. In particular it

was noted that the master-slave scheme used by *Burgee et al.* produced a severe load imbalance among the nodes where the total idle time amounted to one half of the total processing time. In particular *Krasteva et al.* implemented two dynamic balancing strategies so that the load could be effectively redistributed among processors. *Krasteva* and co-workers implemented two dynamic loading algorithms to achieve better speedups on and Intel Paragon: random polling (RP) and global round robin with message combining (GRR-MC). Unlike the strategies used by *Burgee et al.* the two dynamic loading strategies considered by *Krasteva* and co-workers conserved linearity in speedups as more processors were added. In other words using the RP and the GGR-MC strategies *Krasteva* and co-workers were able to halve the processing time when the number of nodes was doubled.

Eldred and Schimel (1999) presented a study where they advocated the use of a mix of coarse-grained and fine-grained parallelism for optimal parallelization. The authors recommend that when using parallel computers coarse-grained parallelism should be used as much as possible since minimal inter-processor communication exists between processors. However, when using massively parallel computers there may be fewer analyses to perform than nodes, or some nodes may be idle for other reasons. In this case it is suggested that a mix of fine-grained and coarse-grained should be used.

Coarse-grain parallelism was also exploited by *Kroo et al.* (1994) in their MDO study of an aircraft using the collaborative optimization method. Collaborative optimization involves the decomposition of the design problem in sub-problems that can then be run in parallel on a parallel machine. Using this strategy a system-level optimizer is used to pass parameters to the different sub-problems. These parameters are adjusted by the system level optimizer to maximize the objective function while satisfying the

system-level constraints. Using coarse-grain parallelization and the collaborative optimization strategy *Kroo* and co-workers were able to decrease the computational time by 69% for the problem they considered.

Numerical Noise

The solution of optimization problems is more difficult in the presence of numerical noise. Noisy behavior of the objective function and of constraints can occur as a result of incomplete convergence of iterative processes, the use of adaptive numerical algorithms, round off errors, discrete representation of continuous physical objects, and discontinuities in the functions that are being calculated. Such noise is typically manifested as a high-frequency, low amplitude variation in the results obtained from computer simulations as the design parameters vary. When numerical optimization is attempted, this oscillatory behavior can create numerous, spurious local optima and can cause slow convergence or even convergence failure. *Giunta et al.* (1994) faced these issues when trying to optimizing the TOGW of an HSCT aircraft due to noise in aerodynamic response. *Giunta et al.* successfully applied response surface to remove noise form the aerodynamic response. Figure 3 shows an example of the noise in the volumetric drag coefficient C_{DWave} as function of the wing.

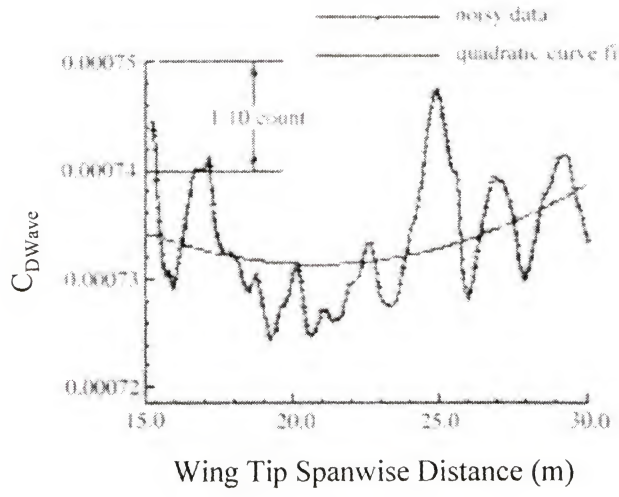
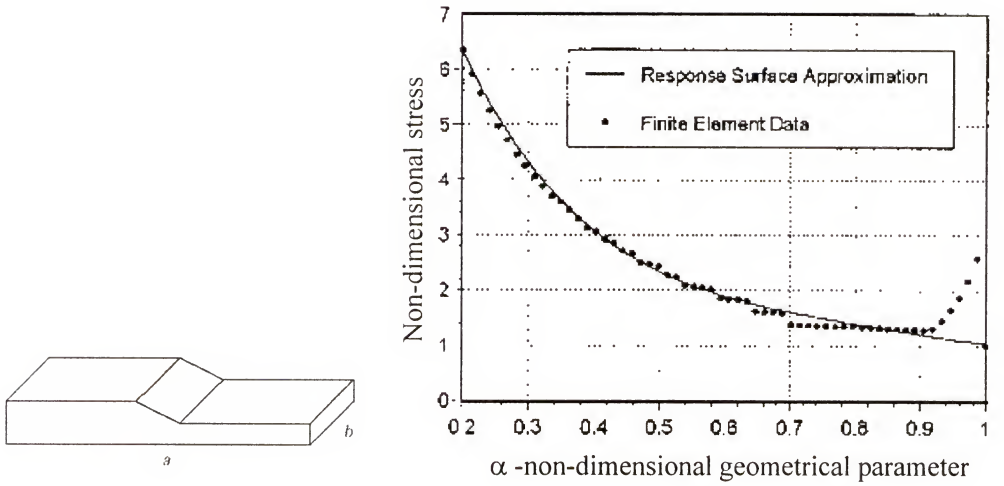


Figure 3: C_{DWave} noisy behavior from *Giunta et al.*

The response surface approximation used and shown in Figure 3 filters out the numerical noise present in the data while preserving the overall trend in the variation in C_{DWave}. Although the C_{DWave} calculations are noisy, the numerical noise has small magnitude and therefore does not affect much the C_{DWave} function values. However it does affect greatly the calculation of derivatives. From Figure 3 it is clear that not only the derivatives of C_{DWave} with respect to the wing tip spanwise distance are inaccurate but that the fluctuation of the derivatives also creates spurious local optima. As a consequence traditional derivative-based optimization algorithms become unreliable.

When a response surface approximation to the C_{DWave} is fitted to the data a smooth response is obtained. Since physically C_{DWave} should change smoothly with the span, the response surface is likely to be more accurate then the actual aerodynamic data. A similar situation where the response surface was probably more accurate than the data used for its construction was encountered by *Venter et al.* (1996) as shown in Figure 4b.



a) b)

Figure 4:Variable thickness plate geometry a) and noise is FE analysis b) as in *Venter et al.*

Venter et al. encountered some inconsistencies in the data obtained from the finite element model (represented by the dots in Figure 4b) in studying a plate with variable thickness. In particular, it yielded less accurate finite element models for plates with values of α close to 1, where α was defined as the ratio of the thickness of the thicker part of the stepped plate to the thinner one (see Figure 4a). Note that for $\alpha = 1.0$ the non-dimensional stress should be equal to one. The inconsistencies were attributed to the mesh generation procedure that did not provide finite element models with a consistent level of fidelity. The results presented in Figure 4b indicate that the response surface approximation eliminated these inconsistencies, and indicate that not all differences between predicted and numerical response values should be considered to be errors.

Free et al. (1987) compared derivative based optimization algorithms and response surface based algorithm. They found that the methods used were comparable in accuracy and computational cost when no noise was present, but that response surface based optimization consistently performed better when noise was present.

The ability of response surface to filter out with noise was exploited by *Etman et al.* (1996) in a crash worthiness optimization problem. Physical crash tests of newly designed vehicles are used to determine the vehicle's crash worthiness. However, these physical tests are very costly, and therefore car producers generally tend to use instead simulation software, such as MADYMO (*TNO Road-Vehicles Research Institute,* 1994) in the design stage. Although crash simulations are computationally very expensive, they are much cheaper than physical tests. As a result, the number of designs that can be analyzed is limited. Designers are urged to find methods that allow optimization using as few crash analyses as possible. *Etman et al.* found that the behavior of injury parameters as a function of the design variables was not smooth. Figure 5 shows the head injury criterion (HIC) as a function of the airbag hole size when all the other parameters are held constant.

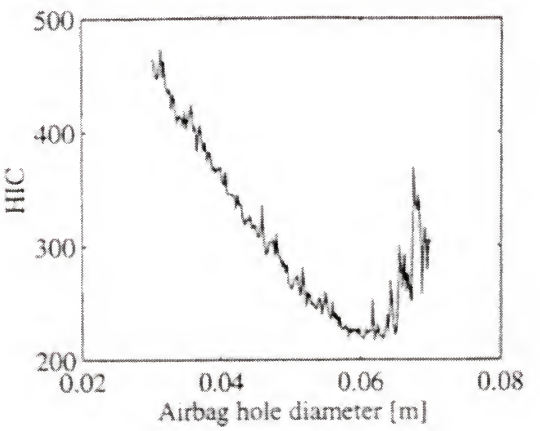
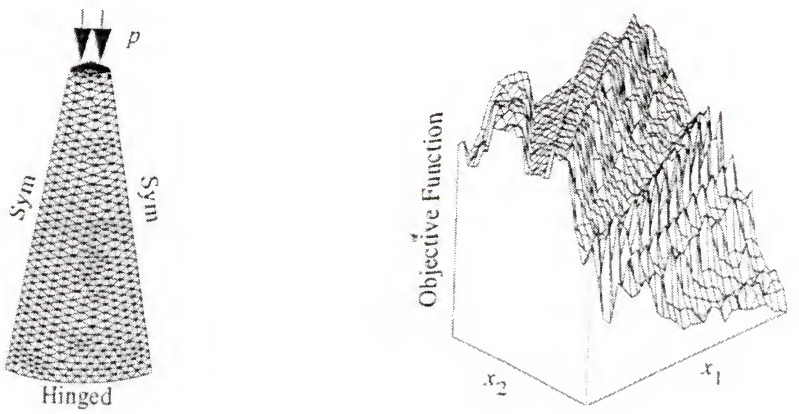


Figure 5: Noisy response as obtained by *Etman et al.* for the head injury criterion (HIC) as a function of the airbag hole diameter

From Figure 5, it is clear that the noise caused gradients to be inaccurate, thus rendering traditional derivative based optimization algorithms inappropriate for this problem. *Etman* and co-workers used an iterative process where sequences of linear

response surface were obtained in sub-domains until convergence to an optimum was achieved. Six design variables were used to optimize the HIC subject to injury constraints on head, chest, and shoulders produced by the airbag and the seat belt. *Schoofs et al.* (1992) approached a similar problem of crash worthiness when optimizing the performance of a child seat as a function of several neck injury parameters. In contrast to *Etman* and co-workers, *Schoofs et al.* used quadratic response surface to approximate the behavior of child seat through the entire design domain. However *Schoofs* and co-workers used only two design variables.

Noisy response functions were also encountered by *Toropov et al.* (1996) in their study of a conical continuous fiber reinforced thermoplastic (CFRTP) material shell under compressive loads. Figure 6a shows the finite element model used by *Toropov* and co-workers, while Figure 6b shows the objective function for the conical shell as a function of height of the shell (x_1) and the base radius (x_2).



a) b)

Figure 6: Finite element model a) and noisy response b) obtained by *Toropov at al.*

Once again it is evident from Figure 6b that gradient based optimization algorithms would fail to reach the optimum. The use of response surface approximations

can smooth out the objective function providing a more accurate prediction than the FE analyses themselves.

Kaufman et al. (1996) encountered significant numerical noise in the optimization of the HSCT wing structural weight material. Variations in the objective function due to numerical noise amounted up to as much as 27%. It is sometimes difficult to separate noise generated by the computer simulations from fitting errors. Apparent errors in the response surface approximations must be carefully interpreted. Indeed a substantial part of the error may represent a desirable smoothing of response surface from the results of the structural optimization. In other words when dealing with noisy functions, response surface can be more accurate than the computer simulation used to produce the data on which the response surface approximation is based on.

Outliers

In general, the accuracy of response surface approximations obtained from fitting experimental data (both numerical and physical) obtainable can be improved if "outliers" can be detected and either excluded or repaired. Outliers are design points whose computed or measured response is far from the predicted response given by the response surface approximation, and their number is usually a small fraction of the number of data points used to construct the response surface approximation. Numerical noise is often present in structural optimization problems as shown in a previous section. However it is often difficult to decide if outliers are due to numerical noise, modeling errors or simply reflect mistakes in the program or data preparation. In general, simply omitting outliers from the regression procedure is not a good practice. Robust regression numerical procedures, such as iteratively reweighted least square (IRLS), can be used to identify outliers in the data set and then once identified the outliers can be more carefully studied,

and the source of the large deviation from the response surface approximation found. In a laboratory the experiments can be repeated or in a computer simulation convergence parameters can be tightened. If the outliers were due to a mistake or convergence failure they can be repaired. On the other hand if the response values remain unaltered even after repair, there is an indication that bias error might be present and that perhaps a higher order polynomial or different shape functions should be considered for the RS.

Papila and Haftka (1999) repaired outliers when fitting response surface approximations to optimize the wing bending material weight of an HSCT aircraft. They fitted a quadratic response surface approximation in five design variables (21 terms) to 43 data points of the wing bending material weight. An IRLS procedure identified four out of the 43 HSCT configurations as outliers. Once the outliers were identified, the cause of poor accuracy was detected as optimization premature convergence. The outlier points where repaired by repeating the optimization with different convergence criteria, leading to more accurate approximation. By using response surface and IRLS to detect outliers and repair them the authors were able to cut by half the predicted errors of the response surface approximation.

Response surface based on gradient information

Response surface techniques have been originally developed for fitting experimentally-hence noisy data. Consequently, these methods historically focused on fitting function data without any attempt to use derivative data, since derivative data is rarely available from physical experiments. In computer simulations, derivatives are often available, frequently at a small computational cost. As a consequence, there has been a growing interest in considering response surface that incorporate both function and derivative data.

There are two cases where using derivative data is particularly attractive. The first case is when the response surface data is generated in the course of an optimization process. Most optimization procedures require derivatives, and so, as the optimization progresses, function and derivative data is accumulated at points previously visited by the optimization procedure. The second case where derivatives are attractive is when the computational cost of a derivative is much lower than the cost of a function evaluation. This happens, for example, when adjoint methods are used to calculate the derivatives of one function with respect to many design variables. *Toomer et al.* (1998) have used such adjoint generated gradients of aerodynamic drag for response surface construction. However, rather than using derivatives directly, they used the derivatives to generate extrapolated function values at nearby points and then used these additional function values with standard response surface methods.

While derivative information is attractive to use when available inexpensively, it is of questionable value when it is much noisier than the function information or when derivatives do not properly exist at many points in the design space. Hence, it is observed that practical use of derivative data for response surface construction is potentially hindered by three complications. First derivatives, like functions, can be noisy and derivatives may be less accurate than the corresponding functions. Second, although the function values might be intrinsically continuous, the derivatives may contain discontinuities. Finally, some derivatives may be calculated inexpensively, while others may be equally expensive as function evaluations. In that case, adding additional function evaluations probably make more sense than adding data obtained by an expensive sensitivity analysis.

In the past, several papers addressed response surface approximation constructed on the basis of both function values and derivatives. There were mainly two motivations to incorporate information on derivatives. First, it has been attempted to establish response surfaces which provide better approximations for the derivatives. Second, reducing computing time and formulation of efficient numerical schemes has been a the other motivation, especially when derivatives can be calculated cheaply.

Sharifzadeh et al. (1989) considered a manufacturing process. The latter is modeled as $y = f(x) + \varepsilon$, here y denotes a random output vector. The random vector ε has zero mean and is mainly determined by manufacturing practice. The input variables $\mathbf{x}$ are given as $\mathbf{x} = x_0 + \eta$, with x_0 a designer specified vector of input variables and η are perturbations independent of ε and with zero mean. *Sharifzadeh* and co-workers put the main focus on the transmitted variance, *var(y)* i.e. variances in the output variables caused by perturbations in the input variables. The transmitted variance can be approximated by $\nabla \mathbf{f}^{\mathrm{T}} \operatorname{var}(\eta) \nabla \mathbf{f}$ It is clearly seen that, this calls for accurate derivatives of the functions, which have been approximated using RS. *Sharifzadeh et al.* report that application of a standard response surface techniques - based on function values only - led to proper results for function values, but inaccurate derivatives. In order to improve derivative accuracy, at each plan point a second point is added whose coordinates are slightly perturbed from the original ones. The perturbation in taken as the vector of standard deviations in the input variables. In summary, the method described by *Sharifzadeh et al.* attempts to improve the accuracy of derivatives by using additional information on directional derivative data, the latter is incorporated by additional function evaluations in the neighborhood of regular points. The authors showed that the methodology applied was successful in predicting function values, but success for the

transmitted variability was more limited. It seemed to be inherently much harder to model variability because it depends on derivatives which tend to be less like polynomials than the functions themselves.

Response surface approximations that are based only on gradients have also been used. *Ho et al.* (1991) proposed a so-called gradient surface methodology and used it for unconstrained optimization. Within the optimization process, linear approximations are built for the gradients of a response function, i.e. approximations to the function itself are not considered. The approximations are determined using a standard least squares fit in terms of the gradients.

Toropov and co-workers (*Malkov and Toropov,* 1991) and *Toropov et al.,* 1993) used mid-range multi-point approximations in their optimization technique. At each cycle of the optimization a number of new function evaluations is carried out. The corresponding design points can be viewed as a plan of experiments in the actual search sub-domain, with its dimension and location controlled by a move limit-strategy. The response surface are fitted with function values using a weighted least square technique, with weights reflecting the relative importance of data to the optimization process. For example, points close to the boundary between feasible and infeasible domains are weighted more heavily compared to the points with severe constraint violation. This procedure attempts to achieve the best response surface approximations at potential optimum regions. If design sensitivities are available, then it is advocated to include this information in the response surface construction at the condition that Design sensitivities should be normalized by the norm of the gradient. The relative influence of design sensitivities is controlled by a user defined non-dimensional parameter. In this optimization strategy, design sensitivities are included to improve overall efficiency.

There are two aspects that may improve efficiency. First, more accurate and complex approximations might be established if sensitivities are available. Second, if design sensitivities may be evaluated cheaply, then the number of design evaluations can be reduced as additional data is provided in terms of design sensitivities. *Toropov* and co-workers compared the results obtained using their method with the results present in the literature for several optimization problems. For the test problems the optimization method including gradient information proposed by *Toropov* and co-workers converged to the optimum solution in fewer cycles then the methods available in the literature. For the test problems the algorithm without gradient information was not used. *Toropov* and co-workers compared the performance of their algorithm with and without gradient information only for one problem. It was found that for that specific problem the algorithm converged in fewer cycles when the gradient information were not included in the approximate function formulation. However such a comparison was carried out creating the same approximating function and using the same number of design points for both the algorithms with and without gradient information. Possibly the comparison is somewhat unfair since the full potential of the use of gradient information was not used since a higher order approximation could have been generated at each cycle when gradient information were used.

Derivatives and function values were included in a regression procedure for global approximation by *Van Campen et al.* (1995) and *Etman* (1992). The derivatives and the functions were combined along by a weighted least squares formulation. It was assumed that the random errors in the function values and derivatives for a single point in the design space are correlated, and that the covariance may vary form point to point. However, it was also assumed that errors for different points are uncorrelated. Weights in

the regression were selected as the inverse of the covariance matrix. The latter in not known a priori and can be approximated once an estimation of the response surface is available, leading to an iterative procedure for determination of regression coefficients, covariances and weights. *Etman* concluded that incorporating derivatives in response surface approximation construction is most beneficial if the approximations offer sufficient flexibility, i.e. higher order approximation functions with relatively large number of regression coefficients are required, otherwise he recommends to use function values only.

Wang and Grandhi (1994) proposed an algorithm for optimal design of frame structures that uses gradient information. An essential component of this algorithm is the construction of a multivariate spline approximation, based on information gained during the optimization process. The coefficients of the multivariate spline approximations are determined from a modified least square formulation. The main motivation for including derivatives in the least square formulation is better accuracy. For the problem that the authors investigated design sensitivities could be easily calculated. A similar procedure was followed for reliability oriented optimization by *Wang and Grandhi* (1995). The method used by the authors can be classified as a mid-range method. The main difference between the work of *Wang and Grandhi* and *Toropov*'s method is that the latter allows additional points to be added in the actual search sub-domain. Moreover, *Wang and Grandhi* adapt the complexity of their spline approximations as more information becomes gradually available. In contrast to this, *Toropov* and co-workers tend to use the same type of response surface approximation for all cycles of the optimization process.

Motivated by gradients that can be evaluated cheaply, *Dyck and Lowther* (1998) investigated response surface techniques using information on both function values and

derivatives. To be more specific, an elaboration on the interpolated moving least square (IMLS) formulation (*Lancaster and Salkauskas,* 1986) has been considered. *Dyck and Lowther* applied their methodology to the designs of loudspeakers magnet assembly. In this formulation the weight allocated to each of the sample points depends on the distance between the point for which the response surface approximation is being evaluated and the sample points. The resulting set of equations is solved using singular value decomposition. It is suggested that weights selected for derivatives should be different form those attributed to function values.

Least square formulation have been used extensively for the formulation of stress recovery techniques. These techniques are used to construct stress distributions that have better accuracy than the underlying finite element solution. Often the main purpose is to generate local error estimates, which can be used as a basis for adaptive mesh refinement. In contrast to most stress recovery schemes, *Riggs et al.* (1997) also included information on derivatives in the least square formulation. In this way, C^1-continuos recovered stress distributions are established. The weighting of derivatives versus function values was done using a user-defined non-dimensional parameter and the length of the element sides. It was reported that the results are remarkably insensitive to the selection of the user-defined parameter. In a more recent work by *Tesser et al.* (1998) the formulation was augmented with an additional term to reduce the "curvature" of the recovered stress distribution. The purpose was to improve the smoothing capabilities of the response surface approximation.

However, current use of gradient information in the construction of response surface approximation is still based on the theory developed for response surface not based on gradient information. The design of experiment techniques and the statistical

tools that are used to reduce the number of terms in a response surface while improving the accuracy of the response surface are not yet available for response surface including derivatives. The proposed research will address these issues.

CHAPTER 3
RESPONSE SURFACE METHODOLOGY

Response Surface methodology can be defined as a collection of statistical and mathematical techniques useful for developing, improving, and optimizing processes. The most extensive application of response surface methodology can be found in the industrial world, in situations where several input variables influence some performance measure, called the response, in a way that is difficult or impossible to describe with a rigorous mathematical formulation. In these situations it might be possible to derive an expression for the performance measure based on the response values obtained from experiments at some particular combination of the input variables. The expression of the performance measure obtained through experiments is called response surface (RS). This chapter introduces the basic mathematical and statistical concepts that belong to the response surface methodology. These concepts allow deriving a response surface approximation from experiment values, choose the number and the location of the experiments in the design space, and measure the accuracy of the response surface once it is obtained.

Fitting an approximation to given data

In this section it is understood that the experimental data are provided to the designer and that his task is to obtain an accurate response surface approximation using the given data. For the moment it is also assumed that the designer is not concerned with the number and the location of the experiments in the design space.

It is assumed that a relationship that describes the performance measure of a phenomenon that it is under investigation exists, and that this relationship is a function of some design variables. It is important to note that this relationship does not have to be known explicitly or even be possible to be known exactly. The response that is to be approximated is denoted as y, and in principle it could represent any measurable quantity such as stresses, chemical concentration, crop production, national debt etc. The true response of the performance measure is denoted by η and it is a function of the design variable vector $\mathbf{x}$ of dimension n_{dv}, that is

$$\eta = \eta(x_1, x_2, .., x_{n_{dv}}) = \eta(\mathbf{x}) \tag{3.1}$$

The response is generally obtained from experiments, in which case η denotes the mean or expected response value. At any combination of the design variables, the value of the experimentally obtained y differs from the expected value η due to random experimental error ε with mean of zero. The relation between y and the vector $\mathbf{x}$ of n_{dv} design variables can be written as

$$y = \eta(\mathbf{x}, \boldsymbol{\beta}) + \varepsilon \tag{3.2}$$

where β is the vector of parameters. If a response surface approximation $\hat{y}$ is used to approximate η the previous equation can be written as

$$\hat{y}(\mathbf{x}) = \eta(\mathbf{x}, \mathbf{b}) \tag{3.3}$$

where $\mathbf{b}$ an estimate of β. The choice of the function $\eta(\mathbf{x},\boldsymbol{\beta})$ is very important in order to achieve a good approximation. However this is also a very difficult task since the true form of $\eta(\mathbf{x},\boldsymbol{\beta})$ is not known, and therefore there is usually no way of knowing which class of function give the best approximation. The form of the approximation $\eta(\mathbf{x},\boldsymbol{\beta})$

depends on shape functions and coefficients of the shape functions. For example one of the most common forms of $\eta(\mathbf{x},\boldsymbol{\beta})$ is

$$\eta(\mathbf{x},\boldsymbol{\beta}) = \sum_{i=1}^{p} \beta_i f_i(x) \tag{3.4}$$

where p is the number of coefficients and shape functions of $\eta(\mathbf{x},\boldsymbol{\beta})$. For example, the yield per acre η of some farmland could be a function of the average seasonal temperature T and the amount of fertilizer F used by a farmer. Forms of η that might approximate the true response as a function of the design variables T and F in a form similar to the form of Eq. (3.4) could be

$$\eta(T,F) = \beta_0 + \beta_1 T + \beta_2 F^2 \tag{3.5}$$

or in a different form

$$\eta(T,F) = \beta_0 + \beta_1 T^{1.2} + \frac{1}{F + \beta_2} \tag{3.6}$$

Both Eq. (3.5) and Eq. (3.6) have three constant terms β_0, β_1 and β_2. In Eq. (3.5) η is a linear function of the coefficients β_i. In contrast the last term of Eq. (3.6) is non-linear in β_2. Note that the linearity or non-linearity refers to the coefficients of the response surface and not the shape functions. Nonlinear forms like Eq. (3.6) lead to so called "non-linear regression " and are not used in this work. The term "linear RS" is confusing because for many people it implies first order polynomials. What is linear is the regression not the response surface form.

Once the form of η is set, that is its shape functions are chosen, the designer has to find the estimates b_i of β_i using the available data. The best set of parameters refers to the parameters that minimize some error measure. The most popular of these measures is

by far the square root of the error mean square ($\sqrt{MSE}$ or *RMSE*), where MSE is defined by

$$MSE = \frac{1}{n-p}\sum_{i=1}^{n_v}\left[y_i - \eta(\mathbf{x}_i, \boldsymbol{\beta})\right]^2 \tag{3.7}$$

where n is the number of experiments available to the designer and p is the number of parameters in the response surface. Other error measures that sometimes are used to find the best set of parameters estimates b_i are the average error (*AE*)

$$AE = \frac{1}{n}\sum_{i=1}^{n_v}\left|y_i - \eta(\mathbf{x}_i, \boldsymbol{\beta})\right| \tag{3.8}$$

and the maximum error (*ME*)

$$ME = \max_{i \in n}\left|y_i - \eta(\mathbf{x}_i, \boldsymbol{\beta})\right| \tag{3.9}$$

In this study, the $\sqrt{MSE}$ or *RMSE* is used as the measure of the error in order to find the best set of estimates b_i to the parameters β_i.

When the only error present in the experimental data is random, the value of the components of the vector of parameters β that would be found if we had an infinite number of experiments is sometimes thought as the "true" value of the parameters of the RS. However, with a finite number of experiments only an estimate of β, herein denoted by **b**, can be obtained. As a general rule of thumb the larger the amount of data available the more accurate the estimates **b,** even though this is not always true.

In order to illustrate how to calculate the estimates **b** that minimize the *RSME* from the experimental data, it is assumed that the response surface approximation has p coefficients or that the vector of coefficients β has dimension p (in the yield example of

Eq. (3.5) and Eq. (3.6) $p = 3$). Moreover it is assumed that there is no modeling error, i.e. η can represent the true response that is

$$y_i = \eta(\mathbf{x}_i, \boldsymbol{\beta}) + \varepsilon_i \quad \text{for } i = 1,.., n \tag{3.10}$$

For the derivation of the **b** vector of estimates parameters in presence of bias error see *Khuri and Cornell* (1996).

Since the response surface approximation is linear in the parameters β_i Eq. (3.4) can be conveniently written in matrix form as

$$\boldsymbol{\eta} = X\boldsymbol{\beta} \tag{3.11}$$

or

$$\begin{Bmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_n \end{Bmatrix} = \begin{bmatrix} f_1(\mathbf{x}_1) & f_2(\mathbf{x}_1) & \cdots & f_p(\mathbf{x}_1) \\ f_1(\mathbf{x}_2) & f_2(\mathbf{x}_2) & \cdots & f_p(\mathbf{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ f_1(\mathbf{x}_n) & f_2(\mathbf{x}_n) & \cdots & f_p(\mathbf{x}_n) \end{bmatrix} \begin{Bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{Bmatrix} \tag{3.12}$$

where **X** is a matrix of response surface shape functions evaluated at the data points x_i, and the vector β contains the unknown coefficients β_i of the response surface approximation η. The shape functions used in response surface methodology are usually low order polynomials, and then Eq (3.12) takes a simple form, since $f_i(\mathbf{x}_j)$ are monomials.

For example, if the response surface approximation is a first order polynomial function as

$$\eta(\mathbf{x}, \boldsymbol{\beta}) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_p x_p \tag{3.13}$$

its matrix form given by n design points is

$$\begin{Bmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_n \end{Bmatrix} = \begin{bmatrix} 1 & \mathbf{x}_{11} & \cdots & \mathbf{x}_{1p} \\ 1 & \mathbf{x}_{21} & \cdots & \mathbf{x}_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \mathbf{x}_{n1} & \cdots & \mathbf{x}_{np} \end{bmatrix} \begin{Bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{Bmatrix} \tag{3.14}$$

where the first index in the elements of the **X** matrix in Eq. (3.14) refers to the data number and the second refers to the design variable. The difference between the model and the j^{th} experiment, r_j called the residual is

$$r_j = y_j - \hat{y}(\boldsymbol{x}_j) \tag{3.15}$$

Equation (3.15) can be written in vector form so to include all the n points as

$$\mathbf{r} = \mathbf{y} - \mathbf{X}\boldsymbol{\beta} \tag{3.16}$$

where **y** is the vector whose elements are the values of the experimental data. The $\sqrt{MSE}$ or *RSME* that follows from Eq. (3.17) can be conveniently written as

$$\sqrt{MSE} = RMSE = \sqrt{\frac{1}{n-p}\mathbf{r}^{\mathrm{T}}\mathbf{r}} \tag{3.17}$$

From Eq. (3.17) is obvious that the vector of estimated parameters **b** that minimizes the *RMSE* is the same one that minimizes $\boldsymbol{r}^{\mathrm{T}}\boldsymbol{r}$. Substituting Eq. (3.16) in $\boldsymbol{r}^{\mathrm{T}}\boldsymbol{r}$ we obtain

$$\mathbf{r}^{\mathrm{T}}\mathbf{r} = \mathbf{y}^{\mathrm{T}}\mathbf{y} - \mathbf{y}^{\mathrm{T}}\mathbf{X}\boldsymbol{\beta} - \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}\mathbf{y} + \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}^{\mathrm{T}}\mathbf{X}\boldsymbol{\beta} \tag{3.18}$$

The vector **b** that minimizes the *RMSE* can be found by differentiating $\boldsymbol{r}^{\mathrm{T}}\boldsymbol{r}$ with respect to the components of β and set them to zero. This leads to

$$\mathbf{X}^{\mathrm{T}}\mathbf{X}\mathbf{b} = \mathbf{X}^{\mathrm{T}}\mathbf{y} \tag{3.19}$$

Equation (3.19) reduces to a p x p system of equations called the *normal equation*, that can be solved for **b** as

$$\mathbf{b} = (\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{y} \tag{3.20}$$

For moderate and large p, $(\mathbf{X}^{\mathrm{T}}\mathbf{X})$ is very ill-conditioned and $\mathbf{b}$ is computed using numerical algorithms such as QR decomposition that solve directly the equation $\mathbf{y} = \mathbf{X}\mathbf{b}$ for the least square solution. Furthermore it is important to note that in order to improve numerical stability is a good practice to scale all variables so that each variable changes in the range (-1,+1).

If the error ε satisfy the conditions of *i*) $E(\varepsilon) = 0$, *ii*) $\mathrm{Var}(\varepsilon) = E(\varepsilon^{\mathrm{T}}\varepsilon) = \sigma^2\mathbf{I}_n$ where E(*) indicates the expected value and $\mathbf{I}_n$ is the identity matrix of order n then the model

$$\boldsymbol{\eta} = \mathbf{X}\boldsymbol{\beta} \tag{3.21}$$

is correct and $\mathbf{b}$ is an unbiased estimator of β (*Khuri and Cornell*, 1996)

$$E(\mathbf{b}) = \boldsymbol{\beta} \tag{3.22}$$

and

$$Var(\mathbf{b}) = (\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\sigma^2 \tag{3.23}$$

It is important to note that σ^2 is rarely known and that an unbiased estimate s^2 of σ^2 is often used

$$\sigma^2 \approx s^2 = MSE = \frac{SSE}{n-p} \tag{3.24}$$

where in Eq. (3.24) the sum of the square errors (SSE) is the sum of squares of errors unaccounted for by the fitted model is

$$SSE = \mathbf{y}^{\mathrm{T}}\mathbf{y} - \mathbf{b}^{\mathrm{T}}\mathbf{X}^{\mathrm{T}}\mathbf{y} \tag{3.25}$$

Once the value of the vector components $\mathbf{b}$ is obtained, it is possible to calculate the variance of the parameter estimates b_i from Eq. (3.23).

Evaluation of predictive capabilities

Response surface approximations are often used to replace constraints or objective functions in optimization problems. Optimization problems however have the general tendency of exploiting weakness in the formulation of the RS. In terms of response surface approximation, this means that the optimizer tends to drive the optimum design to regions where the approximations are inaccurate. Highly accurate response surface approximations are thus a requirement for structural optimization application. Various statistical tools are available in order to evaluate the predictive capabilities of an response surface approximation from the data used to generate them. In addition, in this dissertation the accuracy of response surface approximations is ultimately checked by supplementary sets of experiments. If the prediction capabilities of the response surface are satisfactory, the second set of data can then be added to the first to obtain a more accurate RS.

For the first set of data, the coefficient of determination R^2, the adjusted R^2 statistic, R^2_{Adj}, the percent root mean square error *%RMSE*, and the predicted sum of squares *PRESS* are usually calculated. Moreover, for each coefficient of the response surface approximations the t-statistic is calculated. The t-statistics indicate the probability of a coefficient of being zero rather than the value calculated from the regression procedure. The coefficient of determination R^2 is the proportion of total variation of the values of y_i about the mean $\bar{y}$ explained by the fitted model.

$$R^2 = \frac{SSR}{SSY} = 1 - \frac{SSE}{SSY} \tag{3.26}$$

where, *SSY* is the sum of squares of the deviations of observed y_i about their average value $\bar{y}$

$$SSY = \mathbf{y}^T \mathbf{y} - \frac{(\mathbf{1}^T \mathbf{y})^2}{n} \tag{3.27}$$

where the vector **1** is a vector whose components are all equal to the unity and it has the same dimension of **y,** that is *n.* *SSR* is the sum of squares of the variation of $\hat{y}$ from the same average $\bar{y}$

$$SSR = \mathbf{b}^T \mathbf{X}^T \mathbf{y} - \frac{(\mathbf{1}^T \mathbf{y})^2}{n} \tag{3.28}$$

and SSE is given in Eq. (3.25).

The coefficient of determination R^2 ranges between 1 and 0, with 1 generally indicating a good fit and 0 a bad one. Unfortunately, a value close to 1 does not automatically imply a good response surface approximation because additional parameters added to a response surface approximation will always increase the R^2 value without necessarily increasing the predictive capabilities of the response surface approximation. The adjusted coefficient of determination R^2_{Adj} statistic is an alternative measure of the explained variability that is often used and has the desirable property that its value does not necessarily increase when adding additional parameters to a response surface approximation. When a large difference exists between R^2 and R^2_{Adj}, a good chance exists that the response surface approximation contains insignificant parameters. The R^2_{Adj} statistic is defined as

$$R^2_{Adj} = 1 - \frac{SSE/(n-p)}{SSY/(n-1)} = 1 - \left(\frac{(n-1)}{n-p}\right)\left(1 - R^2\right) \tag{3.29}$$

The percentage error is defined in relation to $\tilde{y}$ the average absolute value of the measured response data, where

$$\tilde{y} = \frac{1}{n}\sum_{i=1}^{n}|y_i| \tag{3.30}$$

so that the *%RMSE* is defined as

$$\%RMSE = \frac{100}{\tilde{y}}\sqrt{\frac{SSE}{\chi}} \tag{3.31}$$

where $\chi = (n\text{-}p)$ for the first set of data used for the fit and $\chi = n$ for the additional set of data used to check the prediction capabilities of the RS. For the first set of data the *(n-p)* factor is used to obtain an unbiased estimator of the mean square error of the response surface approximation, compensating for the fact the error is calculated at the data points used to construct the RS. Even though Eq. (3.31) is based on an unbiased estimator of the mean square error, the *%RMSE* value may be overly optimistic in estimating the predictive capabilities of the RS. An alternative estimate of the predictive capabilities of a response surface is obtained when defining the *%RMSE* based on the *PRESS* statistic. The *PRESS* statistic is calculated by excluding one data point, say data point *i.* The response surface obtained from the remaining *(n-1)* data points is then used to predict the response at the excluded data point, denoted by $\hat{y}_{(i)}$ The prediction error at the excluded data point $r_{(i)}$ may then be written as

$$r_{(i)} = y_i - \hat{y}_{(i)} \tag{3.32}$$

and is referred to as the i^{th} *PRESS* residual. This procedure is then repeated for all data points of the original data set and the resulting *PRESS* residuals are summed to obtain the *PRESS* statistic as follows

$$PRESS = \sum_{i=1}^{n} r_{(i)}^2 \tag{3.33}$$

The i^{th} *PRESS* residual can be calculated from the regression performed using all the n data points as *Myers and Montgomery* (1995)

$$r_{(i)} = \frac{r_i}{1 - h_{ii}} \tag{3.34}$$

where h_{ii} denotes the diagonal terms of the matrix H defined as

$$\mathbf{H} = \mathbf{X}(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}} \tag{3.35}$$

However for large problems Eq. (3.33) may not be accurate enough. The *%RMSE* based on the *PRESS* statistic is then defined as

$$\%RMSE_{PRESS} = \frac{100}{\tilde{y}}\sqrt{\frac{PRESS}{n}} \tag{3.36}$$

Besides the *%RMSE* other error measures associated with the second set of points are sometimes used, and these are the percent average error *%AE* and the percent maximum error *%ME*. These last two quantities are defined as

$$\%AE = \frac{100}{\tilde{y}}\frac{\sum_{i=1}^{n}|y_i - \hat{y}_i|}{n} \tag{3.37}$$

and

$$\%ME = \underset{i \in n}{Max}\left(\frac{100}{\tilde{y}}|y_i - \hat{y}_i|\right) \tag{3.38}$$

Elimination of redundant parameters

Since the exact form of the actual response or mean value η is rarely known, polynomial functions are generally assumed as response surface approximations. The polynomial approximation usually includes redundant parameters, or parameters poorly

characterized by the experiments. These parameters might increase the error in the response surface approximation at points other than the data, thus decreasing their predictive capabilities. Eliminating these parameters from the response surface approximation is an important step in constructing accurate response surface approximations.

Several methods exist for selecting the best subset of parameters from a general polynomial function to be used as a final response surface approximation. Examples include the "all-possible-regression" and the "stepwise regression" procedures (*Myers and Montgomery,* 1995). The all-possible-regression procedure requires that all possible combination of parameters from the general response surface approximation be examined. The best subset of parameters is then chosen based on a suitable criterion. The procedure requires the examination of 2^p partial response surface approximations for p parameters in the general response surface approximation.

The stepwise regression procedure was developed to provide an alternative to the high computational cost associated with the all-possible-regression procedure, especially for response surface approximations with a large number of parameters. The stepwise regression procedure represents a family of procedures that examine only a small number of partial response surface approximations by either adding or deleting one parameter at a time. There are forward, backward, and mixed stepwise regression procedures. The forward stepwise procedure assumes a response surface approximation with only one parameter, the intercept or zero order term. Parameters are then added one at the time by selecting the parameter that results in the largest increase in the predictive capabilities of the current response surface approximation. This process of adding parameters is

repeated until a specified termination criterion is satisfied. In contrast, the backward stepwise regression procedure starts with a response surface approximation that includes all the parameters of the general response surface approximation. From this response surface approximation, one parameter at a time is deleted, selecting the parameter with a minimum influence on the predictive capabilities of the resulting response surface approximation. Finally, the mixed stepwise regression procedure consists of a combination of both the forward and the backward stepwise procedures. In particular this procedure usually starts with an initial response surface approximation that includes all the parameters of the general response surface approximation and then applies backward stepwise regression. At each step of the backward stepwise regression, the least significant parameter is removed form the current response surface approximation. Additionally, at each step all removed parameters are re-examined for possible inclusion in the response surface approximation. The criterion used to determine when to re-include a parameter is usually more stringent than the criterion used to remove a parameter.

It is important to underline that stepwise regression procedures do not guarantee the best subset of parameters and must be used with caution. However since in the JMP program *SAS Institute* (1998) used in this study, the three types of stepwise regression are available, all three were used. In general good results were obtained by applying first the forward, then the backward and finally the mixed stepwise regression and checking that the three methods converged to the same form of the RS. When the three methods did not lead to the same result, the response surface with the best termination criterion was chosen.

In the present study, Mallow's C_p statistic was used as the termination criterion to identify the best reduced response surface approximation from the subset of reduced response surface approximations provided by the stepwise regression procedure. The C_p statistic addresses the prediction capabilities of the response surface approximation by estimating a total error. The total error consists of two parts, one resulting from modeling errors (bias) and the other from noise (variance error). An under specified response surface approximation (not enough parameters and shape functions) increases the part of the total error resulting from bias, while an over specified response surface approximation (too many parameters and shape functions) increases the part of the total error resulting from variance. The C_p statistic is defined as

$$C_p = \left(\frac{SSE_p}{MSE} \right) - (n - 2p) \tag{3.39}$$

where SSE_p is the sum of squares of the n errors for a response surface approximation with p parameters (including the intercept or constant parameter) and MSE is the error mean square of the errors terms obtained from the response surface approximation with all parameters included as.

$$MSE = \frac{SSE}{n - p} \tag{3.40}$$

Based on the C_p statistics, the best response surface approximation in a set of candidate response surface approximations is the response surface with the lowest C_p value, and a good response surface approximation should at least have a C_p value close to p.

Another method to discard unnecessary parameter and so enhance the prediction capabilities of a response surface approximation relies on the R^2_{Adj} and the student

t-statistic test denoted t_{stat}. In general tests of hypothesis concerning the individual parameters in the proposed model are performed by comparing the parameter estimates b_i in the fitted model the their respective estimated standard error $ESE(b_i)$. It can be shown from Eq. (3.23) *Myers and Montgomery* (1995) that

$$ESE(b_i) = \sqrt{Var(b)_{ii}} = \sqrt{MSE}\sqrt{(\mathbf{X}^T\mathbf{X})^{-1}_{ii}} \tag{3.41}$$

so that t_{stat} is defined as

$$t_{stat} = \frac{b_i}{ESE(b_i)} \tag{3.42}$$

The student test statistic t_{stat} gives an idea of how reliable the estimate of a parameter is. The higher the value of t_{stat} the more reliable is the estimate.

Design of experiments

The previous sections dealt with the issue of constructing a response surface approximation for a given set of data. The predictive capabilities of response surface approximations are strongly dependent on the distribution of design points in the design space. Therefore, rather than just choosing designs at random, it is important to apply some sound strategy for selecting a good set of points at which to carry out experiments. The discipline concerned with design or planning of experiments is known as design of experiments (DOE), and a set of selected design points is referred to as an experimental design.

The most extensive body of DOE work has been occupied with the selection of experimental design is regular design domain, i.e. "box-like" or cuboidal domains, where variables are bounded by simple upper and lower limits. This has resulted in various standard layouts for experimental design, which are often preferred by experimenters due

to their simplicity. However, when the objective is to build a response surface approximation for design optimization, the feasible domain will not typically be a regular one, since constraints tend to produce odd shaped design domains. It is important to note that even though the constraints break the regularity of the design domain they serve the great purpose of considerably reducing the size of the design domain and therefore helping to obtain more accurate response surface approximations *Giunta et al.* (1994).

Factorial designs

In box-like design spaces the range of a design variable between the upper and lower limit is typically divided into levels and experiments are run only at the particular levels of the design variables. A factorial design contains all the combination of the different levels of all the design variables also commonly called factors. It is easy to understand how the factorial design is not affordable for problems that have more than a few design variables and few levels: the number of experiments simply becomes too large. The number of experiments required by a factorial design is $(l)^{n_{dv}}$ where n_{dv} is the number of design variables and l is the number of levels per design variable. For example a factorial design of 7 design variables in 3 levels has 2187 experiments. This is usually too many to be run. However response surface approximations obtained from factorial design are generally accurate because they avoid extrapolation.

There are several special cases of the factorial design that are important to discuss because they are widely used and because they form the basis of other designs of considerable practical value. The most important of these is perhaps the $2^{n_{dv}}$ factorial design. Figure 7 shows a two level factorial design in three design variables that is a 2^3

factorial design. Note that in the figure location of experiments are denoted by a black dots.

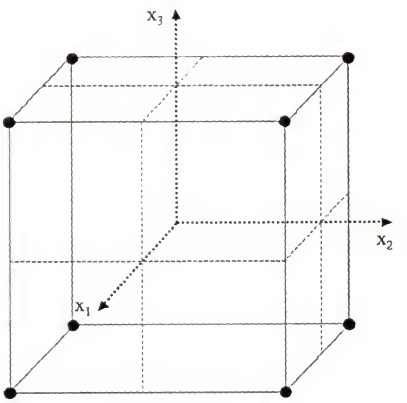


Figure 7: Full factorial design, 2^3

This design is particularly useful in the early stages of experimental work when only the different impact of the design variables on the results is of interest. The design points of a $2^{n_{dv}}$ factorial design in three-dimensional space, i.e. a problem with only three design variables, are the vertices of a cube as shown in Figure 7. Because there are only two levels per design variable only a linear response surface or at most the mixed terms of a quadratic response surface approximation can be obtained from the $2^{n_{dv}}$ factorial design.

The $2^{n_{dv}}$ factorial design has the very attractive properties of *orthogonality* and *rotatability*. Orthogonality and rotatability can be mathematically defined based on the values of the moments of the experimental design. *Khuri and Cornell* (1996) devoted great attention to these two concepts and the reader can refer to that text book. For the moment is enough to say that an orthogonal design is one such that the $(\mathbf{X}^{\mathrm{T}}\mathbf{X})$ design matrix in the normal equation Eq. (3.19) is diagonal which implies that each parameter b_i

in Eq. (3.19) can be estimated independently. A rotatable design instead is one such that the prediction variance $v(\mathbf{x})$ which is given by

$$v(\mathbf{x}) = \sigma^2[f^T(\mathbf{x})\,(\mathbf{X}^T\mathbf{X})^{-1}\,f(\mathbf{x})] \tag{3.43}$$

and it depends only the distance of $\mathbf{x}$ from the center of the design space.

One of the drawbacks of the $2^{n_{dv}}$ factorial design is that it must be assumed that the response surface approximation is linear or at most inclusive of the mixed terms. This limitation inhibits the use of $2^{n_{dv}}$ factorial designs for many problems. When the true function that is being approximated by the response surface approximation is expected to of have considerable curvature, designers often turn to $3^{n_{dv}}$ factorial design. The $3^{n_{dv}}$ factorial designs allow to estimate all the terms of a quadratic response surface approximation. An example of a $3^{n_{dv}}$ factorial design is given in Figure 8 for a three variable case. The 3^3 design shown in Figure 8 is orthogonal but it is not rotatable.

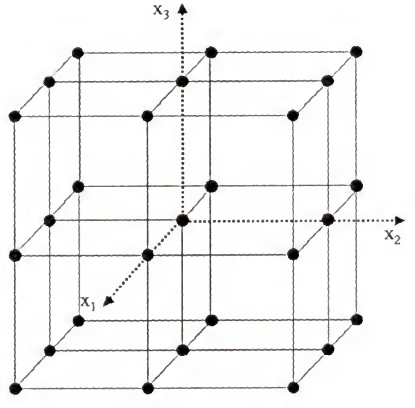


Figure 8: Factorial design 3^3

Fractional factorial designs

Factorial designs unfortunately are affected by the curse of dimensionality, and cannot be used for high dimensions. As an example consider the case of a problem where

the behavior of a physical quantity is described by 10 design variables. Even the $2^{n_{dv}}$ factorial design could hardly be used as 1024 designs would be required and 1024 experiments are too many for most problems. Also note that a linear response surface in 10 design variables has only 11 parameters that have to be estimated and therefore in principle 11 experimental design points would be enough to solve the normal equation. These considerations are even truer for $3^{n_{dv}}$ factorial designs. A three levels factorial design in just seven design variables consists of 2,187 experiments, which is enough to rule out most practical applications. In situations where the number of design variables or the number of levels for design variables is too high designers usually rely on using only a fraction of the design points that constitute the factorial designs.

Designs that use only a fraction of the factorial design points are called fractional factorial design. A fraction of a $2^{n_{dv}}$ design consisting of only $2^{n_{dv}-m}$ treatments combination where *m* is a non-negative integer such that $2^{n_{dv}-m} > n_{dv}+1$ points is called 2^{-m} fraction of a $2^{n_{dv}}$ factorial design. For example, a 2^{-1} fractional design contains one-half the number of points of a $2^{n_{dv}}$ design. The same notation and definitions are used for fractional factorial design with more than two levels. Figure 9 shows a $3^{(3-2)}$ fractional factorial design.

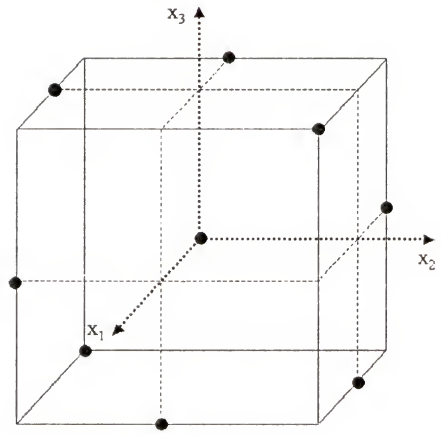


Figure 9: Fractional factorial design, $3^{(3-2)}$

Fractional factorial designs were first introduced by *Finney* (1945). The fraction factorial design used for the estimation of the parameters in the normal equation Eq. (3.19) can be chosen is such a way that it retains some of the properties of the full factorial design, i.e. orthogonaliy and rotatability. It has to be pointed out though that whenever we rely on a fractional factorial design we lose the desirable and important property of interpolation and unavoidably for some portions of the design domain the response surface approximation will predict based on extrapolation which usually leads to poorer accuracy.

Central Composite Designs

The central composite design (CCD) combines full or fractional factorial design with additional points to allow fitting of full quadratic polynomials and it was introduced by Box and Wilson (*Box and Wilson,* 1951). It has become since probably the most popular class of experimental designs for second-order models. The CCD is an augmentation of the two-level full factorial design, with n_c replication of the nominal design (center-point of domain) and the $2n_{dv}$ points obtained by changing one design

variable at a time by a value of $\pm \alpha$. The most used values of α are either 1.0 or $\sqrt[4]{F}$ where F is the number of factorial points. The later value of α one is used in situation where rotatability is desirable. Figure 10 shows an example of CCD for three dimensional problem.

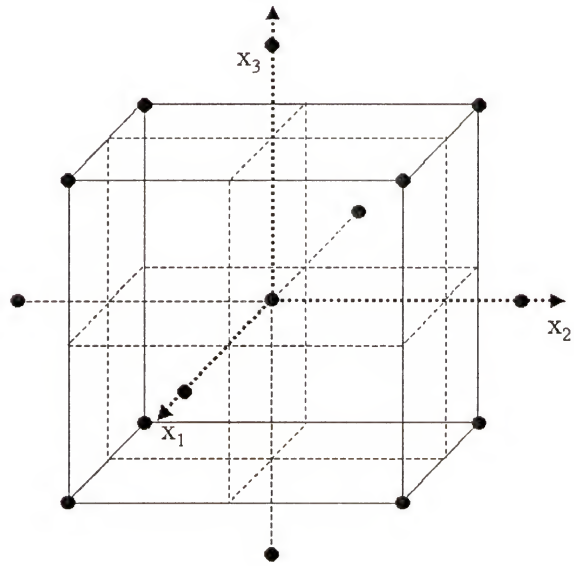


Figure 10: Central composite design in 3 design variables

The nominal design replication serve to improve stability, i.e to make the model less sensitive to errors in data throughout the domain. Of course when the experiments involved are numerical simulations, as it is in the case of this work, there is little sense in repeating the computation several times with the same input values, as the computer will always give the same result. Instead it may be useful to repeat the simulations at slightly perturbed values. For quadratic response models, CCD is an attractive alternative to the full factorial three-level design, since its performance is nearly as good at a much lower cost. In fact, the number of points required totals ($2^{n_{dv}} + 2n_{dv} + n_c$), which is comparable to that of full two-level factorials. For example, a CCD with seven variables and no center

replications ($n_c = 1$) contains 143 sampling points, which is substantially more tractable than the 2187 run required in the corresponding three-level factorial. It is worth noting that some CCD are based on fractional factorial vertices to that the total amount of design points needed are ($2^{n_{dv}-m} + 2n_{dv} + n_c$).

Plackett-Burman designs

The Plackett-Burman (PB) (*Plackett and Burman,* 1946) is yet an other design for first-order response surface approximations. This is a two-level factorial design in n_{dv} variables where the number of design points n is equal to $n_{dv}+1$. These designs are available only when n is a multiple of 4. The objective of this design is to obtain designs that can estimate all main effects with maximum precision possible for $n = n_{dv}+1$.

To construct a Plackett-Burman design in n_{dv} variables, a first row consisting of elements equal to +1 and –1 has to be selected such that the number of positive ones is $\frac{n_{dv}+1}{2}$ and the number of negative ones is $\frac{n_{dv}-1}{2}$. Manipulating the first row cyclically the PB design can be obtained for any problem that satisfies the condition that the number of design variables is a multiple of four. Fortunately most statistical programs as JMP can generate the Plackett-Burman design automatically for the user.

Box-Behnken design

The Box-Behnken design (BBD) *Box and Behnken* (1960) is a three-level incomplete factorial design for the estimation of the parameters in a second-order model. By definition, a three-level incomplete factorial design is a subset of the factorial combinations from the $3^{n_{dv}}$ factorial design. The particular arrangement of the BBD allows the number of design points to increase at the same rate as the number of

polynomial coefficients. The BBD are found by perturbing combinations of two variables in turn, and adding to this n_c replicates of the nominal design. When the experiments are numerical in nature, instead of repeating the nominal design experiment slight perturbations are needed. The two variables are varied, from their nominal value to a low and high level respectively, in all the four possible ways. Thus, since there are $\frac{n_{dv}(n_{dv}+1)}{2}$ ways to select two variables out of n_{dv} ones a BBD consists of totally $2n_{dv}(n_{dv}-1)+n_c+1$ points. For large number of variables, the amount of points asymptotically approaches four times the number of terms to fit in a quadratic polynomial. A seven variable BBD with no center-point replication holds 85 designs, nearly half of the 143 runs required by the CCD and only a fraction of the 2187 needed for a full 3 level factorial design. An example of BBD for three design variables is shown in Figure 11

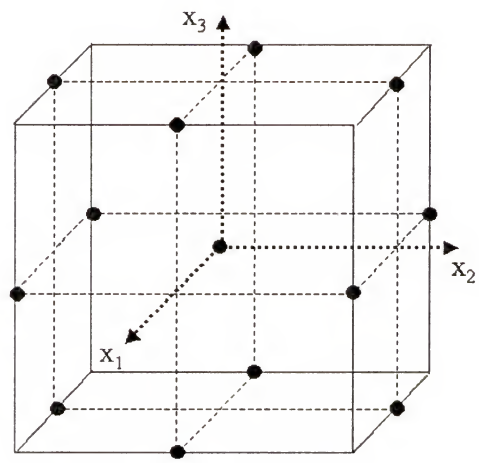


Figure 11: Box-Benhken Design

The inevitable drawback of a BBD, as it can be seen from Fig 1.5 is that corner-regions of the domain are poorly represented. Ideally speaking, as already

mentioned previously, response surface approximations should only be used for predicting responses from interpolation and not extrapolation.

Simplex design

The simplex design is an orthogonal design consisting of $n = n_{dv} + 1$ points where n_{dv} is the number of variables in the first-order model. Unlike the PB design, which has the same number of design points, the simplex can be used for any number of design variables. The design points of the simplex design are located at the vertices of a n_{dv}-dimensional regular sided figure, or a simplex, and are characterized by the fact that the angle, θ, which any two points make with the origin is such that $\cos(\theta) = -\frac{1}{n_{dv}}$ (*Khuri and Cornell,* 1996). The simplex is the design that requires the minimum amount of points to evaluate a first-order RS. A simplex design in three dimensions is shown in Figure 12.

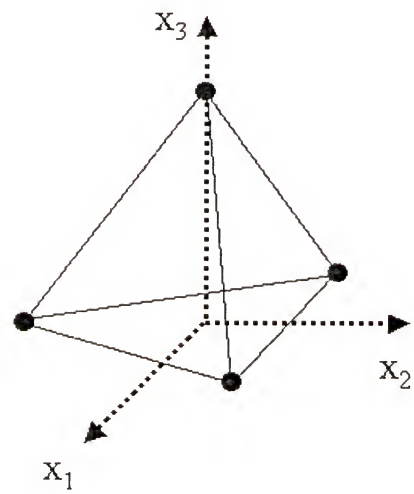


Figure 12: Simplex design

However the main drawbacks of this design is that as the number of design variables increases the volume of the simplex increases exponentially therefore affecting negatively on the accuracy of the RS. Moreover the simplex might have just too few design points to obtain a meaningful response surface approximation. The number of design points of the simplex design is barely enough to perform a fit so that there is no information left to estimate errors and accuracy of the response surface obtained utilizing this design.

D-Optimality

The experimental designs that were presented in the previous sections were all based on the assumption that the experimental design space is a regular space design, either box-like or a simplex. However it is not uncommon for the constraints bounding the design space to create a non-regular type of design space. As mentioned before it is generally advisable to invoke as many constraints as possible to reduce the size of the feasible design space, because this will typically improve the response surface accuracy. Design optimality criteria are usually used for non-regular design spaces and became popular because they are implemented in all major statistical packages of which JMP is no exception. The methods available to generate, evaluate, and compare different design were the results of the work of *Kiefer* (*Kiefer,* 1959 and *Kiefer,* 1961) and *Kiefer and Wolfowitz* (*Kiefer and Wolfowitz,* 1959). With the increasing popularity of computers, algorithms were developed to choose the best design by a software package such as JMP, based on the practitioner's choice of sampling size, model, ranges on variables and other constraints. The most common of the optimality conditions are: D-Optimality, A-Optimality and G-optimality often referred to as alphabetic optimality criteria.

D-Optimality is based on the moment matrix **M**

$$\mathbf{M} = \frac{(\mathbf{X}^{\mathrm{T}}\mathbf{X})}{N} \tag{3.44}$$

The inverse of the moment matrix, $\mathbf{M}^{-1}$, as can be seen from Eq. (3.23), contains the variances and covariances of the regression coefficients, scaled by N/σ^2. As a result, control of the moment matrix implies control of the variances and covariances.

The determinant

$$|\mathbf{M}| = \frac{|\mathbf{X}^{\mathrm{T}}\mathbf{X}|}{N^{p}} \tag{3.45}$$

is inversely proportional to the square of the volume of the confidence region on the regression coefficients under the assumption of independent normal errors with constant variance. The volume of the confidence region reflects how well the set of coefficients are estimated. A small $|\mathbf{X}^{\mathrm{T}}\mathbf{X}|$ and hence a large $|(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}|$ implies a poor estimate of the β vector of coefficients in the model.

A D-optimal design is one in which the relationship of Eq. (3.42) is maximized; that is

$$\underset{\zeta}{Max}\,|\mathbf{M}(\zeta)| \tag{3.46}$$

where *Max* implies that the maximum is taken over all designs. As a result, it is possible to define the D-Efficiency of a particular design ζ^* as

$$D_{eff} = \left(\frac{|\mathbf{M}(\zeta^*)|}{\underset{\zeta}{Max}|\mathbf{M}(\zeta)|} \right)^{1/p} \tag{3.47}$$

The *1/p* power accounts for the *p* parameter estimates being assessed when one computes the determinant of the scaled variance-covariance matrix $(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}$. The

definition of D-efficiency is particularly useful because in principle it allows for comparing designs that have different sample sizes and it can compare different designs in design regions that are not regular or box-type. Besides the advantages just mentioned, the D-optimality does not restrict the number of experiments to be made. Finally, this variance based criteria takes full account of the approximation model, which must be specified prior to the selection of the sampling points, in contrast with the standard experimental designs. For instance, the D-Optimal design in a given design domain will not be identical for a first and a second order response surface approximation. D-Optimal designs tend to select design points at the boundary of the design domain so to increase the design moment **M**.

In principle, the construction of D-optimal designs is an optimization problem, which is generally rather difficult to handle due to the large number of involved variables, ($n \times n_{dv}$) and the complicated nature of the objective function $|\mathbf{X}^T\mathbf{X}|$, and because of the existence of several local minima.

Special algorithms that treat the problems as combinatorial ones have been developed but the resulting integer programming is still a very large one. Solutions algorithms using heuristic techniques must usually settle for suboptimal (yet very good) sets of designs. For example, there are approximately 10^{10} possible ways to pick 10 points out of 50 candidates. A popular way to seek D-Optimal designs are exchange algorithms, which successively remove and add points starting from an initial arbitrarily chosen design (*Fedrov,* 1972 and *Mitchell,* 1974). Also, genetic algorithms have been successfully applied for determining D-optimal designs, see for example *Giunta et al.*

(1994). In this work the DETMAX algorithm first presented by *Mitchell* (1974) as implemented in the commercial statistical package JMP has been used.

Latin Hypercube Sampling

The latin hypercube sampling (LHS) is yet an other design scheme that was successfully used to select design points where to perform experiments to construct response surface approximation. The LHS scheme presented here follows the explanations given by *McKay et al.* (1979) in their study.

Consider a quantity y that is a function of the x_i variables with i =1,..,k. We are interested in constructing an approximation to the function y using n numerical experiment obtained from n design points. The LHS scheme can be used to select the n designs points where to conduct the n numerical experiments. According to the LHS scheme, the range of each design variable is divided into n nonoverlapping intervals on the basis of equal probability. One value from each interval is then selected at random with respect to the probability density in the interval. The n values thus obtained for x_1 are paired in a random manner with the n values of x_2. These n pairs are combined in a random manner with the n values of x_3 to form n triplets, and so on, until n k-tuplets are formed. In these way n design points are generated where the function y can be evaluated.

Consider a simple example where it is desired to generate a LHS of size n = 5 with two input variables. Let us assume that the first random variable x_1 has a normal distribution with a mean value μ and variance σ^2. The five intervals of the variable x_1 are shown in Figure 13.

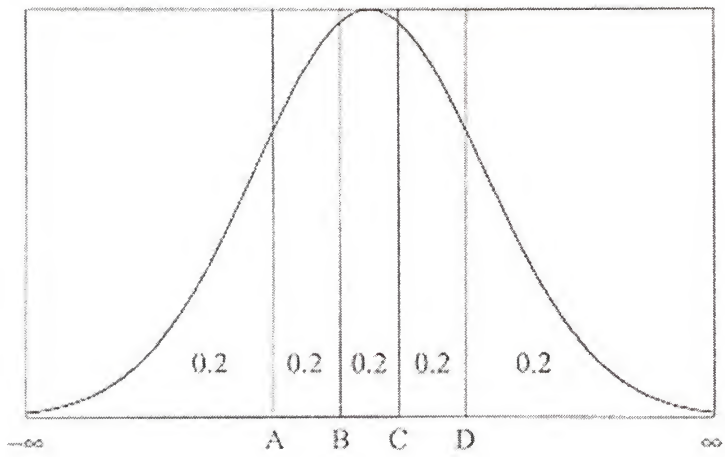


Figure 13: Intervals used with a latin hypercube sample (LHS) of size $n = 5$ in terms of the density function for the normal random variable x_1

The intervals of Figure 13 satisfy the condition that each interval has the same probability, $P(-\infty \le x_1 \le A) = P(A \le x_1 \le B) = \ldots = 0.2$.

It is also assumed for this example that x_2 has a uniform distribution on the interval G to L as shown in Figure 14. The corresponding intervals used in the LHS for x_2 are given in Figure 14.

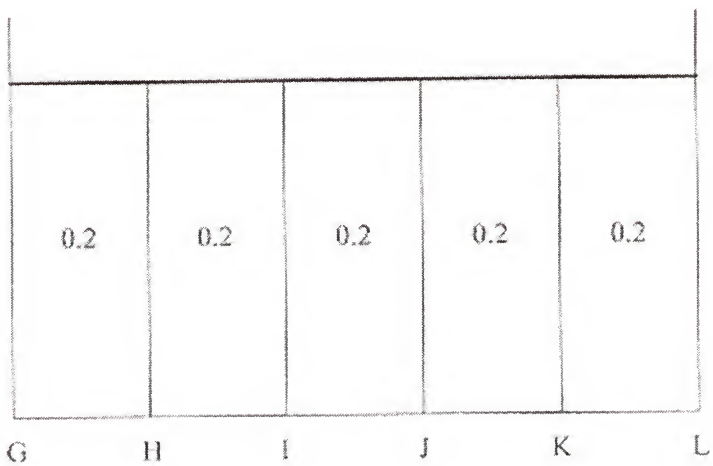


Figure 14: Intervals used with a latin hypercube sample (LHS) of size $n = 5$ in terms of the density function for the uniform random variable x_2

The next step in obtaining the LHS is to pick specific values x_1 and x_2 in each of their five respective intervals. This selection must be done in a random manner with respect to the density in each interval; that is, the selection must reflect the height of the density across the interval. For example in Figure 13 in the $(-\infty,A)$ interval for x_1, values close to A will have a higher probability of selection than those values in the tail of the distribution that extends to $-\infty$. Next the selected values of x_1 and x_2 are randomly paired to form the five required two-dimensional input vectors. For example two random permutations of the integers (1,2,3,4,5) are:

- Permutation Set 1: (3,1,5,2,4)
- Permutation Set 2: (2,4,1,3,5)

By using the respective position within these permutation sets as interval numbers for x_1 (Set 1) and x_2 (Set 2) it is possible to pair the two design variables levels. Thus for the first of n computer runs, the input vector is formed by selecting the specific value of x_1 from the interval number 3 (B to C in Figure 13) and pairing this value with the specific value of x_2 selected from interval number 2 (H to I in Figure 14). The vectors for the second and subsequent runs are then constructed in a similar manner. Once the specific values of each variable are obtained to form the five input vectors, a two-dimensional representation of the LHS such as that given in Figure 15 can be made.

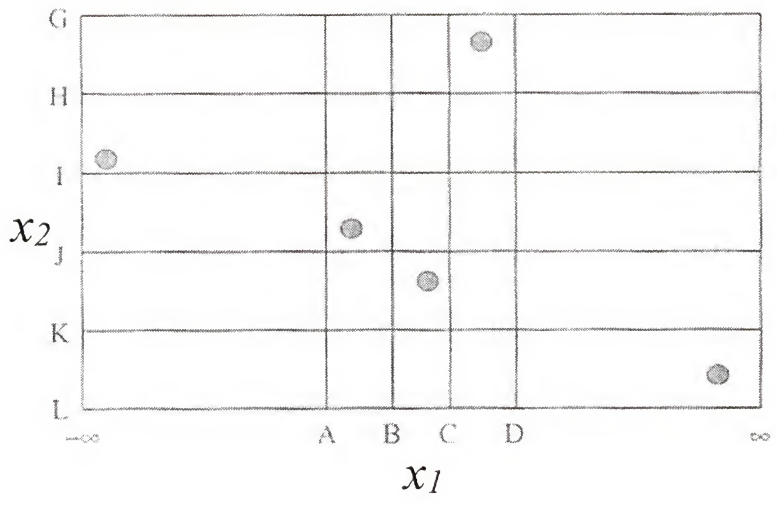


Figure 15: Two dimensional representation of one possible latin hypercube sample (LHS) of size $n = 5$ using one normally distributed design variable x_1, and one uniformly distributed design variable x_2

Note that in Figure 15 all of the intervals for x_1 have been sampled, and the same is true of x_2. In general, a set of n LHS points in k-dimensional space contains one point in each of the intervals for each of the k variables. This feature is particularly important when fitting higher than second order polynomial response surface approximations.

If for example a quartic polynomial in two design variables is desired at least 15 design points are needed since a quartic polynomial has 15 coefficients. However the 15 design points must have at least 5 different levels for each design variables. With less than five levels the coefficients of the fourth power can not be estimated. The LHS guarantees that there are a sufficient number of levels to fit the quartic polynomial if 15 points are chosen.

Other methods for obtaining designs

There are several other properties of the moment matrix **M** that have been proposed as criteria for the purpose of constructing optimal experimental designs. A-Optimality is a criteria that seeks to minimize the trace of the moment matrix **M**. An other method that has been proposed as a valid criterion to evaluate the design efficiency is the G-optimality. The G-Optimality seeks to minimize the maximum prediction variance over the experimental region

$$v(\mathbf{x}) = \sigma^2[f^T(\mathbf{x})\,(\mathbf{X}^T\mathbf{X})^{-1}\,f(\mathbf{x})] \tag{3.48}$$

where $f(\mathbf{x})$ is the function containing the monomials or shape functions of the response surface approximation form.

All the experimental designs previously considered are based on the assumption that the form of the response surface approximation is correct and therefore they are concerned only with reducing the influence of the noise in the data or in other words they are minimum variance designs.

Another class of designs is primarily concerned with modeling or bias errors, and therefore they try to space the design points in the design domain so to minimize the bias error. One drawback of minimum bias designs is that in order to choose the appropriate design that minimizes the bias error, the true function has to be known, a situation that is rarely encountered. In general since polynomials are usually used by designers as approximating functions it is assumed that the true function is a degree higher than the polynomial used as response surface approximation. For example it is possible to find design points that minimize the bias error using for example a quadratic response surface approximation when the true function is a cubic polynomial. *Box and Draper* (1987),

Myers and Montgomery (1995) and *Khuri and Cornell* (1996) discuss minimum bias design in some detail.

Optimization

One of the most attractive features of using response surface methodology in optimization is that due to the least square procedure and smooth shape functions, the resulting approximate response function is smooth. Numerical noise in simulations creates spurious local minima that can easily trap the optimization algorithm and make it difficult to find the best design *Giunta et al.* (1994) and *Venter et al.* (1996).

When the noisy response function obtained from numerical simulation is replaced by a smooth response surface approximation one effectively filters out the numerical noise, thus allowing the use of derivative based optimization algorithms. Derivative based algorithms are efficient algorithms and are readily available commercially. In this dissertation, when not stated otherwise, the generalized reduced gradient algorithm provided with the EXCEL (*Microsoft Corporation, 1998*) program was used to solve the optimization problems. The Excel program provided an easy to use and universally available interactive environment for the solution of optimization problems and for this reason it was preferred to other software.

CHAPTER 4
STRUCTURAL OPTIMIZATION OF A HAT STIFFENER PANEL

Summary

This chapter describes a design study for structural optimization of a hat-stiffened laminated composite panel concept for an upper cover panel of a typical passenger bay of a blended-wing-body (BWB) type transport airplane. The feasibility of a hat-stiffened composite panel concept for an upper cover panel is investigated, and is compared on the basis of weight, to a thick sandwich concept, which has been selected by designers as the baseline concept for the upper cover panel. The upper cover panel concept is designed for two load cases, internal pressure only, and combined internal pressure plus spanwise compression due to wing bending. The structural optimization problem is formulated using the panel weight as the objective function, with constraints on stress and buckling. The spacing of the hat-stiffeners, the thickness of the skin, and the thickness of the components of the hat-stiffener are used as design variables. The initial geometry of the hat-stiffened panel design is determined using the PANDA2 program, by restricting the design to have uniform cross-section in the spanwise direction. Because of the pressure loading, a more efficient design has variable cross-section. Such designs are obtained by combining the STAGS *Rankin et al.* (1998) finite element program with the optimization program in the Microsoft EXCEL (*Microsoft Corporation*, 1998) spreadsheet program using response surfaces. Buckling and stress response surfaces are constructed from multiple STAGS analyses and are used as constraints in the

optimization. The optimization conducted with the response surfaces results in considerable weight savings compared to the uniform cross-section design, albeit in a more complex design. Initial optimization cycles identify a design space where simple approximate analyses, such as Euler-Bernoulli beam theory and Kirchhoff plate theory applied to laminated composites, can be used to predict the behavior of the structure.

Introduction

Major air carriers have expressed the need for larger airplanes to meet the growing demands for air travel, especially in the Pacific Rim and on trans-Atlantic routes between major airports in the USA and in Europe. A blended-wing-body (BWB) 800-passenger aircraft is one of several configurations currently being considered for satisfying this need. As the name implies, a principal feature of a BWB transport is a wide double-deck centerbody that is blended into the wing. Due to the shape of the airplane, the pressurized centerbody region, which includes both the passenger area and the cargo area, is non-circular (see Figure 16).

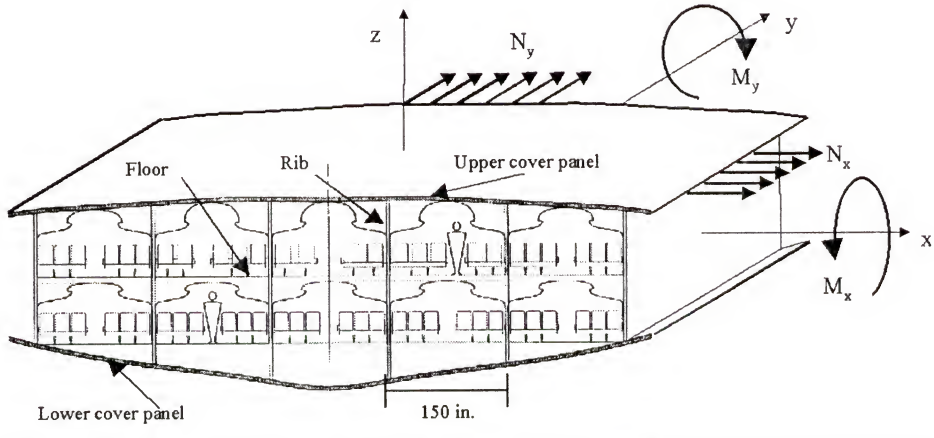


Figure 16: Center-body region of a blended-wing-body transport airplane

The non-circular centerbody region is challenging from the standpoint of structural design, since the upper and lower cover panels carry the overall aircraft body moments M_x, internal pressure, and wing bending moments M_y as shown in Figure 16. To reduce instability problems in the upper cover panel, which is subjected to compressive running loads, N_x, due to wing bending, and to reduce bending stresses due to the internal pressure loading, the inboard centerbody section is divided into several passenger and cargo bays which are separated by chordwise ribs. A typical passenger bay is 150 inches wide, as shown in Figure 16.

A variety of structural configurations for the pressurized upper cover panel structure were considered as part of a BWB design study, led by the McDonnell Douglas Corporation (now The Boeing Company, Long Beach) in collaboration with NASA and four universities (University of Florida, Stanford University, University of Southern California, and Clark Atlanta University). The configurations that were considered in the study include conventional skin-stringer constructions, deep sandwich structure, and separate structures for carrying the internal pressure and wing bending loads. A deep sandwich configuration was chosen as the baseline upper cover panel concept, on the basis of low weight combined with ease of fabrication and assembly. However, some unresolved issues of sandwich structures might prevent their use as primary structure for a BWB airplane. The present study describes a structural optimization study of one of the alternative structural concepts being considered for the upper cover panel of a typical passenger bay of a BWB transport airplane. Specifically, results are presented for a composite hat-stiffened skin panel configuration. The skin and all of the components of the hat-stiffener are built up from graphite-epoxy, warp-knit preforms. Each preform is a

stack of material equivalent to seven layers of unidirectional prepreg with 0, 45 and 90 degree fibers, and has a cured thickness of 0.055 inches. The loading cases used in this preliminary study did not include M_x shown in Figure 16. The wing bending moment M_y was transformed into equivalent N_x compressive and tensile running loads in the upper and lower cover panels respectively. The weight of the upper cover panel obtained for the hat-stiffened panel concept was compared to the thick sandwich concept for the load cases available for this study. A hat-stiffened panel design was considered to be a valid alternative to the sandwich panel design if the weight of the hat-stiffened panel was equal or less than 3.0 lb/ft^2.

The structural optimization problem is formulated using the panel weight as the objective function, with stress and buckling constraints. The spacing of the hat-stiffeners, the thickness of the skin, and the thickness of the components of the hat-stiffener are used as design variables. The thickness variables are integer multiples of the thickness of the basic composite material stack described above. Therefore, the optimization problem is discrete. Furthermore, the structural analyses required to evaluate the constraints are performed by an analysis code that does not have optimization capabilities and is difficult to interface with an optimization program. For such situations, response surface techniques, which create simple approximations of structural responses, have been shown to be useful *Mason et al.* (1994). Response surface techniques also permit simple analysis models to be integrated with more complex analysis models. In the present work, both simple and complex analysis models are integrated in response surfaces that are used in the structural optimization.

Problem Description

The structural configuration considered in the optimization problem is a hat-stiffened skin, upper cover panel of a typical passenger bay, as shown in Figure 17.

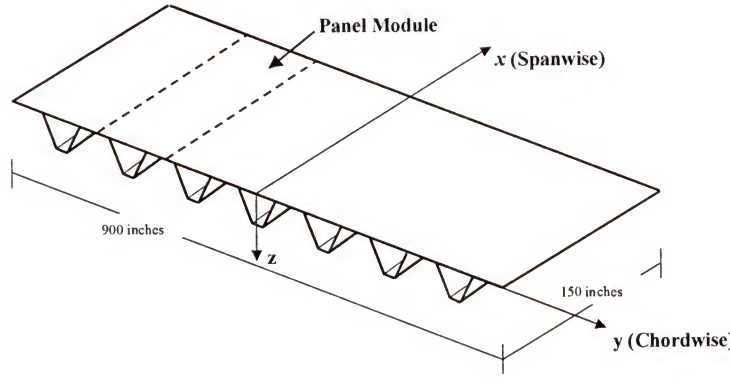


Figure 17: Hat-stiffened skin upper cover panel of the BWB

The panel is 150 inches long in the spanwise (x) direction, and 900 inches long in the chordwise (y) direction. Two loading conditions are considered in the design. The first loading condition is combined internal pressure and spanwise (x) compression. The second loading condition is internal pressure-only. The two load cases considered do not have a load component in the chordwise direction but they were considered to be sufficient for obtaining a comparison of preliminary design concepts for the upper cover panel. The ends of the stiffened panel at $x = 0$ inches, and at $x = 150$ inches are clamped. The clamped boundary condition is achieved by two simple support boundary conditions applied to the base and to the crown of the stiffener. The end of the panel at $x = 0$ inches is restrained from movement in the x-direction while the opposite end of the panel at $x = 150$ inches is constrained to have a uniform u displacement. The unloaded edges of the panel at $y = 0$ inches and at $y = 900$ inches are simply supported with the v displacement component unconstrained.

The skin and the individual components of the hat-stiffener are constructed from graphite-epoxy warp-knit preforms, which were developed under several NASA contracts (*NASA Contracts NAS1 18862*). Each preform is a stack of material equivalent to seven layers of unidirectional prepeg with 0, 45 and 90 degree fibers. The nominal stacking sequence of the preforms used in the skin and in all of the components of the hat-stiffener is [45/-45/0/90/0/-45/45]. Each preform, or stack, has a cured thickness of 0.055 inches. Nominal material properties for a cured stack and the stress allowables that are used in the designs are provided in Table 1.

Table 1: Material properties for graphite-epoxy preforms

E_{11}(msi)	9.25
E_{22}(msi)	4.67
G_{12}(msi)	2.27
ν_{12}	0.397
ε_{all}	$5.40\ 10^{-3}$
σ_{all}(ksi)	50
τ_{all}(ksi)	18
ρ (lb/in^3)	0.0057

Structural Optimization

An initial optimum structural design for the upper cover panel was obtained using the PANDA2 program *Bushnell* (1987). A refined optimum structural design was then obtained by an optimization using response surface techniques. The analyses required to construct the response surfaces were performed using both simple and complex analysis models. The complex structural analyses were performed using the STAGS (STructural Analysis of General Shells) finite element program. The initial design was obtained for a combined load case internal pressure $p = 14.84\ psi$, and spanwise compression $N_x = -4319\ lb/in.$ The optimization for the initial load case is described herein in the

sections "Design Using PANDA2 for Initial Load Condition" and "Design Using Response Surface Techniques for Initial Load Condition". As the airplane design evolved, the design loads for the combined load case were updated to internal pressure $p = 15.59\ psi$ and spanwise compression $N_x = -2879\ lb/in$, and an additional load case of internal pressure-only $p = 18.56\ psi$ was added. The optimum structural design obtained for the initial load case was used as the starting point in the optimization problem for the updated load conditions. The design for the updated load conditions is described herein in the following section "Design for Updated Load Conditions." A flow chart of the history of the different hat-stiffener cover panel structures obtained during the optimization is shown in Figure 18.

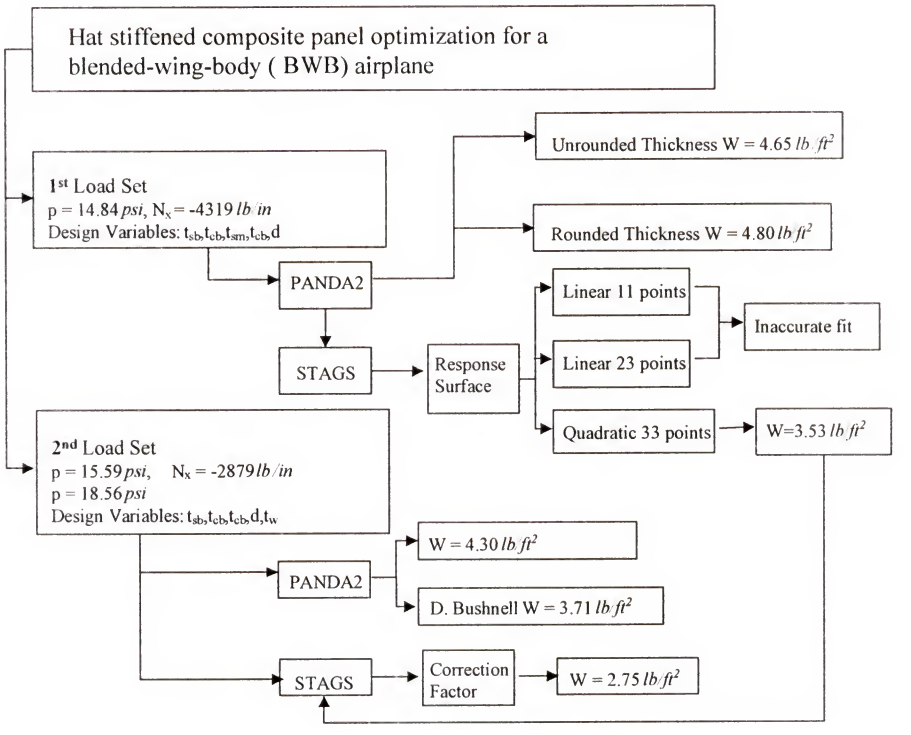


Figure 18: Flow chart of the optimization history of the hat stiffened panel

Design Using PANDA2 for Initial Load Condition

The first step in the structural optimization was to use the PANDA2 program to obtain an initial design. PANDA2 was developed specifically to find minimum weight designs of stiffened, flat or curved panels, or complete cylindrical shells made of laminated composite materials. The stiffeners may run in one, or in two orthogonal directions, and all stiffeners in one direction are assumed to be identical and uniformly spaced. The panels can be loaded by combinations of in-plane loads, edge moments, normal pressure and temperature. Constraints on the design include global and local buckling, stiffener crippling, maximum displacement due to internal pressure loading, maximum tensile or compressive stress along the fiber and normal to the fibers in each lamina, and maximum in-plane shear stress in each lamina. PANDA2 calculates local and general buckling load factors using simple closed form expressions and inexpensive finite difference analyses of discretized models of panel cross-sections. Local buckling is predicted from the analysis of a single panel module, which is assumed to repeat several times over the panel width. A single panel module consists of one stiffener plus the panel skin, with width equal to the spacing between stiffeners. For a typical stiffener PANDA2 finds the axial location in which each element of the panel module has the highest compressive stresses, and performs a stability analysis based on these values. The results of the stability analysis are then used to size each element of the panel module.

In the present problem, the stiffeners run in the spanwise (x) direction, and the panel consists of repetitive modules in the chordwise (y) direction (Figure 17). Furthermore, the properties of the panel are assumed to be uniform in the spanwise (x) direction. A cross-section of a panel module and the design variables used in the PANDA2 optimization are shown in Figure 19.

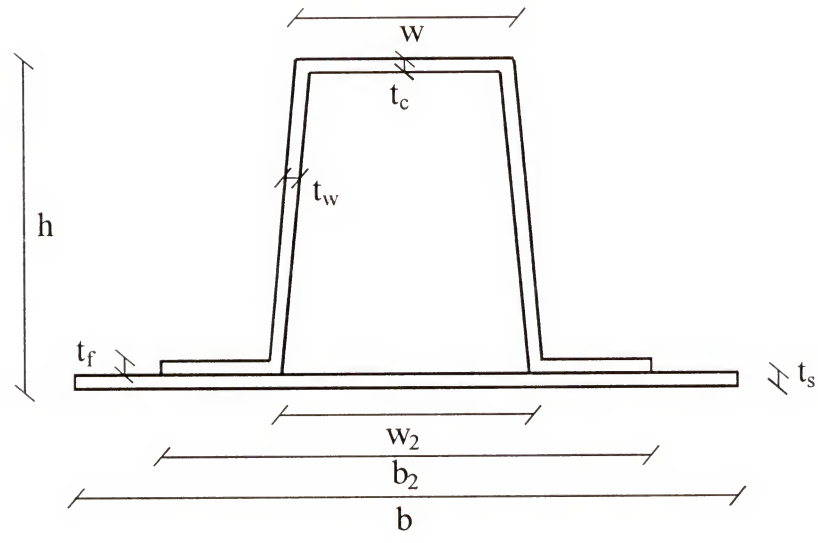


Figure 19: Cross-section of a simple module and design variables used for the PANDA2 optimization

The objective function, the design variables and the side constraints used in the PANDA2 optimization are listed in Table 2.

Table 2: PANDA2 optimization problem (except for weight dimensions are in inches)

Objective function : min {weight}	
Design Variables*: b, b_2, w, w_2, h, t_s, t_f, t_w, t_c	
Side Constraints:	
$6 \leq b \leq 24$	(1)
$b_2 - 0.75b \geq 0$	(2)
$4.5 \leq b_2 \leq 18$	(3)
$2 \leq h \leq 6.5$	(4)
$b_2 - w_2 \geq 2.6$	(5)
$2.9 \leq w_2 \leq 15.4$	(6)
$2.25 \leq w \leq 11.75$	(7)
$0.11 \leq t \leq 1.1$	(8)
Buckling and stresses constraints in PANDA2	(9)

*See Figure 19

The first constraint shown in Table 2 ensures that there are enough equally spaced stiffeners in the panel for a single module model to give a good approximation to the

local skin buckling mode. The second and third constraints are recommended by the PANDA2 user's guide to guarantee numerical stability in the solution procedure. The fourth constraint controls the height of the hat, where the upper value reflects a manufacturing limit. The fifth and sixth constraints ensure that the flange is at least 1.3 inches wide, again for manufacturing reasons, and the seventh constraint ensures that the hat-stiffeners have reasonable proportions. Finally, the eighth constraint sets the upper and lower thickness bounds of the panel skin and of all elements of the hat-stiffener. Additional constraints on the design included global and local buckling, crippling, stiffener pop-off, maximum stresses along and normal to the fibers in each lamina, and maximum in-plane shear stresses within each element of the stiffener. In evaluating the constraints, a factor of safety equal to 1.3 was applied to the stresses and a factor of safety equal to 1.15 was applied to the buckling load factor. The buckling load safety factor reflects a design requirement that the structure will not buckle between design limit load and design ultimate load. The optimum design for the hat-stiffened upper cover panel obtained by the PANDA2 program is provided in the second column in Table 3.

Table 3: PANDA2 original and rounded optimum design

Variable	Original	Rounded
Weight (lb/ft^2)	4.65	4.86
b (in)	13.83	13.83
b_2 (in)	8.4	8.4
h (in)	6.5	6.5
w (in)	4.3	4.3
w_2 (in)	5.8	5.8
t_s (in)	0.22	0.22
t_f(in)	0.34	0.33
t_w(in)	0.20	0.22
t_c(in)	0.33	0.33

As shown in Table 3, the optimum design obtained by the PANDA2 program weighed *4.65 lb/ft*^{2}. The optimum design obtained by PANDA2 features continuous design variables. However, the thickness of the panel skin and the thickness of the individual components of the hat-stiffener are discrete variables, since the thicknesses are limited to integer multiples of the thickness of one warp-knit preform. Therefore, the optimum design obtained by PANDA2 must be rounded to the next feasible value of the discrete design variables. Rounding off the thicknesses to the nearest discrete thicknesses increased the weight of the cover panel to *4.86 lb/ft*^{2}. The rounded-off design is shown in the third column of Table 3. For the rounded-off design shown, the active constraints at the mid-length of the panel were fiber compression in the crown, web buckling, and axial (spanwise direction) strain in the crown. At the ends of the panel, the active constraints included local buckling of the panel skin between the hat-stiffeners, local buckling of the panel skin under the hats, and spanwise compression in the panel skin.

Design Using Response Surface Techniques for Initial Load Condition

For the loading and boundary conditions of the current problem, the design obtained with PANDA2 is conservative. The conservative design results from the imposed requirement that the panel properties are constant in the spanwise direction, while the loading for the present problem is variable in the spanwise direction.

A more efficient design for the stiffened panel was obtained by allowing the cross-section of the panel to vary in the spanwise direction. Based upon the results of the PANDA2 analysis, the panel was divided into three sections: two identical sections at the ends of the panel, and a section in the interior of the panel, as shown in Figure 20.

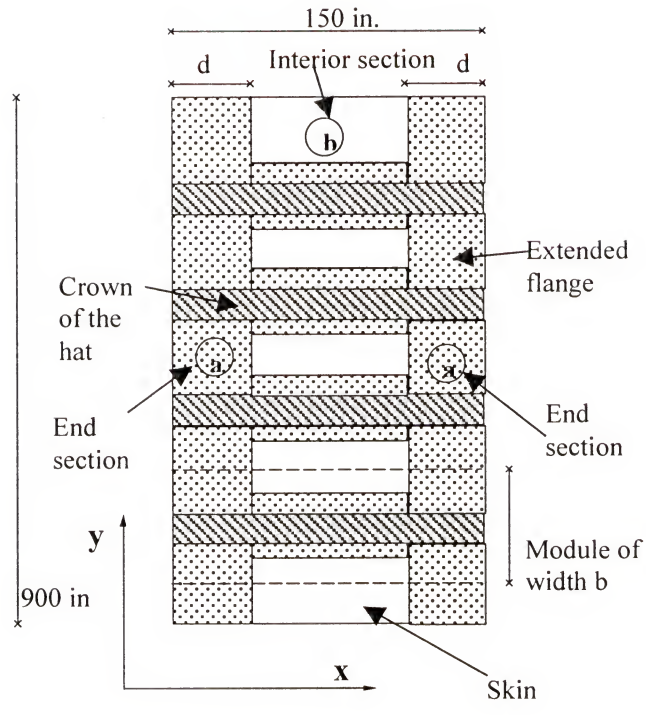
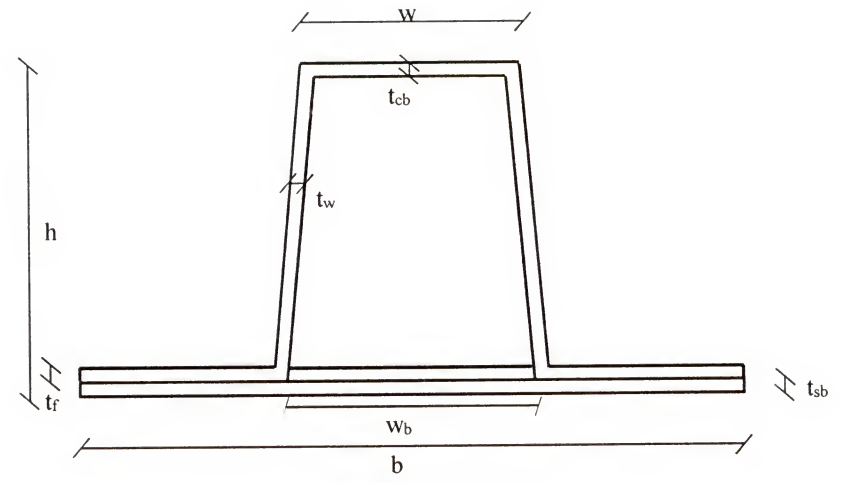
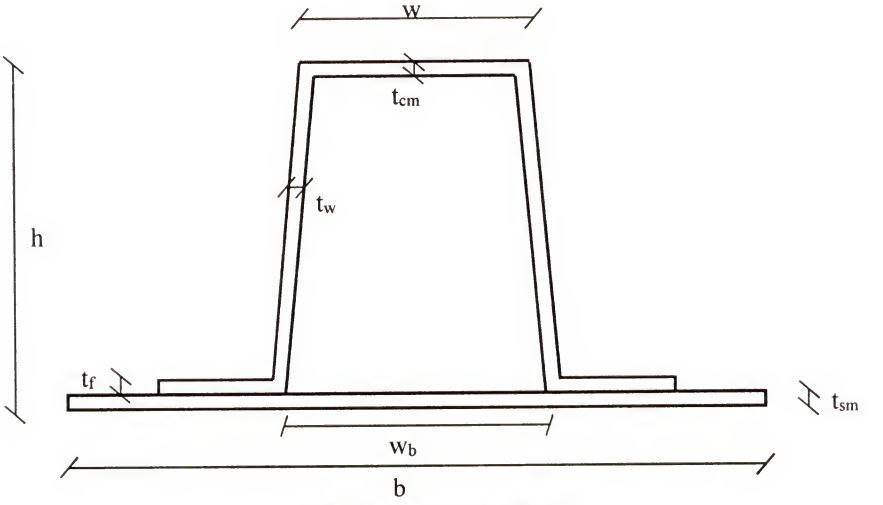


Figure 20: Division of the hat-stiffened panel into three spanwise sections

The tendency of the skin to buckle near the ends of the panel, which was an active constraint in the PANDA2 analysis, was reduced by adding a layer of material equal to the thickness of the flange to the panel skin between the hat stiffeners and to the panel skin under the hat stiffeners. Furthermore, the thickness of the crown of the hat stiffeners in the end sections of the panel was allowed to be different than the thickness of the crown of the hat stiffeners in the interior region of the panel. The cross-sections of a panel module in the end regions of the panel and in the interior region of the panel are shown in Figure 21, respectively.



(a) End section



(b) Interior section

Figure 21: Cross-sectional geometry of a variable thickness panel

The STAGS (*Rankin et al.,* 1998) program was used to obtain accurate buckling loads and stress predictions for the variable cross-section panel shown in Figure 20 and Figure 21. STAGS is a finite element code for the non-linear analysis of general shell structures of arbitrary shape and complexity. STAGS, however, does not have an optimization capability. Response surface techniques, which create simple approximations of structural response, have been shown to be useful for design

optimization when the design variables are discrete and when it is difficult to integrate the analysis and optimization codes. In the current study, STAGS finite element analyses were used to construct response surface models to approximate the panel buckling response and to approximate the mid-surface compressive, spanwise, skin stress at the thickness discontinuity between the end and the interior regions of the panel. There was very little difference between the stresses on the outer surface of the skin and the inner surface of the skin, at the discontinuity between the end and the interior regions of the panel therefore the mid-surface stresses were compared with the maximum stress allowable. For the first loading condition considered in this study, the thickness discontinuity was the only region in the hat stiffened panel where the stress constraint was violated. The response surface approximations created were then used as constraint functions in an optimization procedure to minimize the panel weight. The optimization was performed using a spreadsheet in the *Microsoft EXCEL* (1998) software program that allows discrete design variables. The optimization module in EXCEL, called "SOLVER" is available in several other spreadsheet programs. The design variables used in the optimization are given in Table 4. The thickness of the flange (t_f = 0.11 in.) and the thickness of the web (t_w = 0.22 in.) were kept constant during the optimization process.

Table 4: Design variables used in response surface approximation optimization

Variable	Variable meaning
d	Distance from panel ends to thickness discontinuity
t_{sb}	Skin Thickness near panel ends
t_{sm}	Skin Thickness in the interior of the panel
t_{cb}	Crown thickness the near panel ends
t_{cm}	Crown thickness in the interior of the panel

Both the buckling response surface and the stress response surface were approximated using polynomials. A least squares estimate of the unknown coefficients in the polynomials was obtained by evaluating the structural response (panel buckling load or stress at the thickness discontinuity) at a number of design points that exceeded the number of coefficients in the approximating polynomials. The structural response at the individual design points was determined using the STAGS finite element code. The method of analysis used to perform stress and buckling analysis was linear. In particular the buckling analysis was conducted by performing a linear bifurfaction buckling analysis from a linear prestress state. The internal pressure p and the in-plane load N_x were included in the same load set so that the same buckling load factor was applied to both loads. In all of the STAGS analyses, an additional factor of safety equal to 1.25 was applied to the end load so that a spanwise compression load, $N_x = -5398\ lb./in$ was used.

In order to keep the cost of constructing a response surface as low as possible, while ensuring the accuracy of the response surface model, care must be taken in selecting the design points. Several methods are available for selecting the design points so that the error in the approximation is minimized. In the present study, the D-Optimality criterion, as implemented in the JMP program (*SAS Institute*, 1998) was used

to select the design points. The implementation in the JMP program finds a D-Optimal set of points from a given set of candidate design points in the design domain. The 9604 candidate design points that were used for constructing the response surfaces are defined in Table 5.

Table 5: Candidate points for response surface construction

Design Variable	Minimum (in)	Step (in)	Maximum (in)	Nominal Design (in)
d	12	12	48	36
t_{sb}	0.110	0.055	0.440	0.165
t_{sm}	0.110	0.055	0.440	0.110
t_{cb}	0.110	0.055	0.440	0.220
t_{cm}	0.110	0.055	0.440	0.165

Also given in Table 5 is a nominal design, selected by engineering judgment. The nominal design weighted *3.741 lb/ft*^{2}, had a buckling load factor applied to both the internal pressure and compressive running loads of 1.898, and has spanwise, compressive mid-surface skin stress at the thickness discontinuity equal to 20,288 psi.

Before proceeding with the selection of the design points, the size of the design domain was reduced from 9604 to 740 feasible design points by introducing the following constraints based on the expectations for the optimum design:

- The weight was required to be between 3.0 lb/ft^2 and 4.3 lb/ft^2.
- The skin near the ends of the panel was required to be thicker than the skin in the interior of the panel ($t_{sb} \geq t_{sm}$).
- The crown of the hat-stiffener in the interior of the panel was required to be thicker than the crown in the end section of the panel ($t_{cm} \geq t_{cb}$).

Buckling Response Surface Approximation

The STAGS model used in the analysis of the hat-stiffened panel had approximately 70,000 degrees of freedom. The panel skin and all of the elements of the stiffeners were modeled as branched shells. One linear stress and linear bifurcation buckling analysis using this model required approximately 6,500 CPU seconds on a DEC ALPHA 200 4/166 work station.

Because of the time required to perform each of the STAGS structural analyses, a simple linear polynomial response surface approximation for the buckling load factor was initially constructed first using 11 design points (design points 1-11 in Table 6). Ten of the design points in Table 6 (design points 2-11) were selected using the D-Optimality criterion. Design point 1, the nominal design given in Table 5, was selected by engineering judgement.

Table 6: Structural designs used for the linear response surface

Point #	d (in)	t_{sb} (in)	t_{sm} (in)	t_{cm} (in)	t_{cb} (in)	weight (lb/ft.)	λ	σ_{max} (psi)
1	36	0.165	0.110	0.220	0.165	3.741	1.898	20288
2	12	0.110	0.110	0.110	0.110	3.110	0.554	43031
3	12	0.110	0.110	0.440	0.110	3.747	0.531	42969
4	12	0.220	0.220	0.220	0.220	4.237	2.262	23780
5	12	0.440	0.110	0.440	0.110	4.166	0.821	25978
6	12	0.440	0.110	0.440	0.440	4.288	0.987	27846
7	12	0.440	0.220	0.110	0.110	4.264	0.991	23330
8	48	0.110	0.110	0.440	0.440	4.206	1.232	-
9	48	0.165	0.165	0.440	0.110	4.158	1.207	14295
10	48	0.275	0.110	0.110	0.110	4.288	1.010	15164
11	12	0.165	0.165	0.220	0.165	3.965	1.296	-
12	48	0.220	0.055	0.275	0.220	4.153	0.727	27328
13	48	0.220	0.055	0.275	0.110	3.991	0.749	24871
14	48	0.165	0.110	0.165	0.220	3.314	2.020	17250
15	48	0.110	0.165	0.165	0.220	3.816	1.058	12850
16	48	0.220	0.165	0.165	0.110	4.215	1.460	17514
17	24	0.220	0.165	0.275	0.220	4.142	2.492	21976
18	24	0.220	0.165	0.165	0.110	3.890	2.070	21408
19	24	0.220	0.055	0.165	0.110	3.294	0.201	39861
20	24	0.110	0.165	0.275	0.110	3.781	1.084	20679
21	24	0.110	0.055	0.275	0.110	3.185	0.230	39997
22	24	0.110	0.055	0.165	0.220	3.095	0.205	44642
23	24	0.110	0.055	0.165	0.110	3.014	0.196	44023
24	48	0.110	0.055	0.165	0.110	3.339	1.131	27607
25	24	0.165	0.055	0.165	0.110	3.154	0.199	39004
26	48	0.220	0.165	0.220	0.110	4.261	1.328	12139
27	48	0.110	0.110	0.165	0.110	3.497	1.079	18363
28	24	0.220	0.110	0.165	0.110	3.592	1.058	23939
29	24	0.165	0.055	0.165	0.110	3.154	0.199	-
30	48	0.220	0.165	0.220	0.110	4.261	1.328	-
31	48	0.110	0.110	0.165	0.11	3.497	1.079	-
32	24	0.220	0.110	0.165	0.11	3.592	1.058	-
33	36	0.165	0.110	0.165	0.11	3.612	1.065	-

The linear fit, constructed using these 11 points spanning through the entire design domain, for predicting the buckling load factor was very poor with an estimate of the root mean square error $RMSE = 0.61$, and $R^2_{Adj} = 0.046$. Since the buckling load

factor is approximately equal to 1.0, this *RMSE* error is very large. The parameter R^2_{Adj} indicates the quality of a fit. A perfect fit is indicated by $R^2_{Adj} = 1.0$, and $R^2_{Adj} = 0.0$ represents a very poor fit.

In an attempt to improve the accuracy of the linear response surface approximation, a new design domain was defined around the nominal design. The thickness variables were permitted to change by $\pm$0.055 inches from the nominal design and the length of the end section *d*, was permitted to change by $\pm$12 inches. Thus, each variable could have 3 possible values and a total of 243 design points was generated (81 of the generated points violated the t_{sm} thickness constraint since the nominal design was at the lower bound for this design variable). Twelve design points were selected out of these 243 points using the D-Optimality criterion (points 12-23 in Table 6).

The linear response surface for the buckling load factor obtained using the 12 new points had an $RMSE = 0.54$ and $R^2_{Adj} = 0.56$, indicating that the fit was still unsatisfactory. Moreover, the t_{stat} as defined in Eq. (3.42) of the coefficients were very small. The t_{stat} is a parameter that indicates the confidence in the values of the coefficients obtained and it is defined as the ratio between the value of a coefficient and its variance. A higher confidence is indicated by a higher value of the t_{stat}, which means that the value of the coefficient is much larger than its variance. Since the linear response surface models were not satisfactory, 10 design points were added to the 23 design points that were used in constructing the two linear response surface models described above. Using all the 33 design points in Table 6, a full quadratic polynomial was fitted over the entire design domain. A full quadratic polynomial in five design variables has 21

coefficients. Terms in the polynomial with a low t_{stat} were discarded as long as R^2_{Adj} kept increasing. The quadratic polynomial retained ten coefficients, and is given by:

$$\begin{aligned}\lambda = & -4.199 + 0.1708\ d + 34.14\ t_{sm} - 0.001362\ d^2 \\ & - 31.98\ t_{sb}^2 - 0.4712\ d\ t_{sm} + 72.30\ t_{sm}\ t_{sb} - 95.94\ t_{sm}^2 \\ & - 0.759\ t_{cm}\ d + 22.43\ t_{cm}\ t_{sb}\end{aligned} \tag{4.1}$$

For the polynomial fit of Eq. (4.1), $RMSE = 0.25$ and $R^2_{Adj} = 0.82$. Furthermore, the lowest t_{stat} was 3.17, which indicates reasonable confidence in the coefficients. The accuracy of the response surface approximation was also checked by constructing several response surfaces using only 32 points, and then comparing the response surface predictions at the point left out with the STAGS analysis predictions at that point (a procedure known as *PRESS*, Eq. (3.33)). Based upon these comparisons, the quadratic response surface approximation for the buckling load factor was expected to have less than 20% error.

Stress Response Surface.

A quadratic response surface approximation for the mid-surface compressive, spanwise stress in the skin, at the change in skin thickness between the end sections of the panel and the interior section of the panel, was also constructed. Stress values at all other locations in the panel, for all of the designs analyzed, did not exceed the maximum stress allowable. The response surface approximation obtained with units of (*ksi*) is given by:

$$\begin{aligned}\sigma = {} & 1.087 - 2.011\ d - 95.15\ t_{sb} - 481.5\ t_{sm} \\ & + 0.01503\, d^2 - 1.141\, d\ t_{sb} + 2.816\ t_{sm} d \\ & + 323.4\ t_{sm}\ t_{sb} + 813.6\ t_{sm}^2\end{aligned} \tag{4.2}$$

with, $R^2_{Adj} = 0.96$ and $RMSE = 1924\ psi$. The t_{stat} for all of the coefficients were higher than 3.33. The accuracy of the response surface approximation was also checked by the *PRESS* procedure, which indicated that a maximum error of 20% is expected when the stress response surface is used to estimate the mid-surface compressive, spanwise stresses in the skin near the thickness discontinuity.

Optimization Using the Response Surfaces Approximations

As indicated above, the response surface approximation for the buckling load factor, λ, and the stress response surface had errors of 20%. Therefore the required buckling load factor was increased by 20% from 1.15 to 1.38, and the safety factor on the compressive stresses was increased by 20% from 1.3 to 1.56.

The optimization problem was formulated as shown in Table 7 and solved using a reduced gradient optimizer available in Microsoft EXCEL.

Table 7: Initial optimization problem using response surfaces

Objective Function:	Min{weight}
Design Variables:	d, t_{sb}, t_{sm}, t_{cm}, t_{cb}
Constraints:	$\lambda > 1.38$
	$\sigma < 32051$ (psi)
	$t_{sb} - t_{sm} > 0$
	$t_{cm} - t_{cb} > 0$
	$24 < d < 48$ (in.)
	$0.110 < t_{sb} < 0.220$ (in)
	$0.110 < t_{sm} < 0.165$ (in)
	$0.165 < t_{cm} < 0.275$ (in)
	$0.110 < t_{cb} < 0.220$ (in)

The optimum design obtained, subject to the constraints specified in Table 7, is presented in Table 8. Analysis of the design in Table 8 with STAGS gave a buckling load factor, $\lambda = 1.379$, and a maximum spanwise, stress, $\sigma = -22{,}450$ *psi*, in the skin at the boundary between the end section and the interior section. The variable cross-section design for the upper cover panel, obtained using the response surface approach has a weight of *3.53 lb/ft*^{2} (Table 8) compared to a weight of *4.86 lb/ft*^{2} (Table 3) for the uniform cross section panel design.

Table 8: Initial optimum design obtained using response surface

Parameter	Optimum
d (in)	30
t_{sb} (in)	0.165
t_{sm} (in)	0.110
t_{cm} (in)	0.165
t_{cb} (in)	0.110
Weight (lb/ft^2)	3.533
λ STAGS	1.379
σ (psi.) STAGS	22,450
λ response surface	1.416
σ (psi.) response surface	23,845

Design for Updated Load Conditions

After the optimization described above was completed, the design loads were updated as the overall airplane design changed. A new load case of internal pressure-only $p = 18.56\ psi$ was added, and, in the combined load case, the pressure loading was changed from *14.84 psi* to *15.59 psi* and the in-plane, spanwise load was changed from *4319 lb/in* to *2879 lb/in*. In addition, the buckling and stress safety factors were reduced to 1.0. The reduction in the buckling safety factor reflected a change in design philosophy, allowing local buckling of the skin between the design limit load and the design ultimate load.

PANDA2 was used to optimize the hat stiffened panel with uniform cross-section for the updated combined load case, and a panel weight of $4.30\ lb/ft^2$ was obtained, indicating that the updated loading condition was less severe than the initial loading condition. Moreover, Dr David Bushnell, the developer of PANDA2, obtained an optimal two cross-section hat stiffener panel using PANDA2. For a transition specified at distance of 24.0 inches, the weight obtained by PANDA2 was $3.71\ lb/ft^2$.

To obtain a variable cross-section design, STAGS was again used. The STAGS finite element model of the hat stiffened panel was simplified to include only one panel module, 13.83 inches wide, with symmetry conditions imposed along the sides parallel to the x-axis. Also, the weight calculations were refined to remove duplications in the previous calculations due to intersecting finite elements.

The design for the updated load conditions was initiated using the optimum design obtained in the design cycle for the initial load condition as a starting point. In the optimization for the initial load condition, the thickness of the skin in the interior section of the panel, t_{sm}, remained at its lower limit throughout the optimization process and was

not an active variable in the design. In addition, for the optimum design presented in Table 8, the webs account for approximately 40% of the total panel weight. Based upon these observations, the design variable list was modified by including the web thickness, t_w, as a design variable and by removing the thickness of the skin in the interior section of the panel, t_{sm}, from the design variable list.

A design domain for the updated optimization was generated around the optimum design point obtained in the previous optimization. The design domain was limited to 162 new design points by introducing the following restrictions. The thickness of the skin in the end sections of the panel and the thickness of the crown in the interior section of the panel, were permitted to change by $\pm$0.055 inches from the previous optimum design (Table 8). The thickness of the crown in the end sections of the panel was allowed to be twice as thick as it was for the initial load case in order to handle the new load case of internal pressure-only that induces high tensile stresses in the crown near the panel ends. The web was constrained to have a thickness equal to 0.165 inches or 0.220 inches. A thinner web would violate the buckling constraint, and a thicker web would result in unacceptably heavy designs. The distance from the panel ends to the thickness discontinuity, d, was permitted to change by $\pm$ 6 inches from the previous optimum design value. The thickness of the skin in the interior section of the panel, t_{sm}, was equal to 0.110 inches; and t_{sm} was held constant during the optimization process. The flange thickness, which was equal to 0.110 inches, was also held constant during the optimization process.

The JMP program was used to select a D-Optimal set of 26 new design points from the 162 candidate design points. STAGS was used to conduct linear analyses of

each of the 26 designs for the internal-pressure-only load case and for the combined load case. Table 9 shows the 26 designs and the results obtained from the STAGS analyses for the buckling load factors λ_p and λ_c for the internal pressure-only load case and for the combined load case, respectively.

Table 9: Structural designs near the initial optimum design that was used for the updated response surface optimization

Point #	d (in)	t_{sb} (in)	t_{cm} (in)	t_{cb} (in)	t_w (in)	Weight (lb/ft^2)	λ_p	λ_c	σ_{Max} (psi)
1	30	0.220	0.110	0.110	0.220	3.505	0.848	0.805	56,858
2	30	0.165	0.165	0.165	0.220	3.410	1.680	1.496	46,161
3	30	0.110	0.110	0.110	0.165	2.755	0.661	0.620	-
4	36	0.220	0.220	0.165	0.165	3.366	2.565	1.756	-
5	36	0.220	0.110	0.165	0.165	3.321	0.800	0.695	-
6	36	0.110	0.220	0.165	0.220	3.341	1.000	0.893	-
7	36	0.110	0.220	0.165	0.165	2.943	0.895	0.779	-
8	36	0.110	0.110	0.165	0.220	3.186	0.865	0.797	-
9	36	0.110	0.110	0.165	0.165	2.786	0.761	0.652	-
10	24	0.220	0.220	0.165	0.220	3.602	2.736	1.206	-
11	24	0.220	0.220	0.165	0.165	3.207	2.516	1.003	-
12	24	0.220	0.110	0.165	0.220	3.399	0.870	0.803	-
13	24	0.220	0.110	0.165	0.165	3.002	0.716	0.657	-
14	24	0.110	0.220	0.165	0.220	3.322	1.005	0.906	-
15	24	0.110	0.110	0.165	0.220	3.120	0.832	0.764	-
16	36	0.220	0.165	0.110	0.165	3.290	1.452	1.238	64,761
17	36	0.220	0.110	0.110	0.165	3.212	0.694	0.664	66,607
18	36	0.165	0.220	0.110	0.220	3.553	1.698	1.508	55,149
19	36	0.110	0.110	0.110	0.220	3.189	0.814	0.769	58,858
20	30	0.165	0.165	0.110	0.220	3.420	1.661	1.484	57,009
21	30	0.110	0.220	0.110	0.165	3.334	0.892	0.890	63,389
22	30	0.220	0.220	0.110	0.220	4.000	2.705	1.260	54,627
23	24	0.165	0.110	0.110	0.165	3.262	0.665	0.623	66,618
24	24	0.110	0.220	0.110	0.165	3.324	0.893	0.785	61,609
25	24	0.110	0.220	0.110	0.165	2.926	0.971	0.874	56,235
26	24	0.110	0.110	0.110	0.165	2.722	0.650	0.607	67,113

The buckling load factors were obtained from a linear bifurcation buckling analysis from a linear prestress state. For the combined load case, the internal pressure

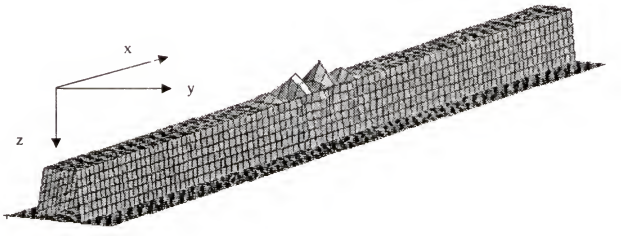
load and spanwise compressive load were included on one load set and the buckling load was applied to the combined load. The pressure-only load generated high tensile stresses in the crown of the hat stiffener due to bending. The tensile mid-surface spanwise stresses in the crown for the internal pressure-only load case, calculated at 3.5 inches from the panel end, are also provided in Table 9. The dashes in Table 9 indicate designs for which stress calculations were not performed.

The stress calculations performed with STAGS indicated that the stress allowable was exceeded in the crown of the hat stiffener, for the internal pressure-only load case only. For this load case, several designs failed due to large, spanwise, tensile stresses in the crown of the hat stiffener near the end sections of the panel. Furthermore, the STAGS analyses showed that for the two load cases considered in the present analysis, a simple beam analysis provides a good approximation of the stress distribution along the panel length. Since the stress calculations performed using the beam approximation are inexpensive, the approximate beam analysis was used to calculate the stresses in the crown of the hat stiffener at 3.5 inches from the panel ends, for all of the 162 structures. The stress results from the approximate beam analysis were compared with the STAGS analysis stress results for 13 structures (points 1-2 and points 16-26 in Table 9. The difference between the simple beam predictions and the STAGS analysis predictions ranged from 0.1% to 6.5%.

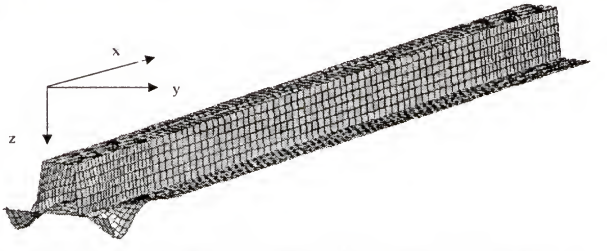
Approximate Buckling Analysis.

The STAGS analyses for the 26 hat-stiffened cover panel design provided in Table 9 showed that for buckling load factors less than 1.0 or close to 1.0, buckling was generally local and confined to a single element of the hat stiffened panel (i.e., only the

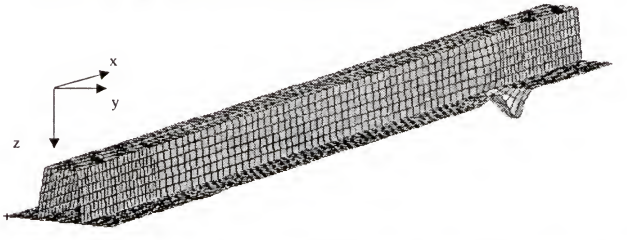
cap of the hat stiffener or the panel skin had any appreciable deformation). This observation is demonstrated in Figure 22, where buckling mode shapes are shown for four design points.



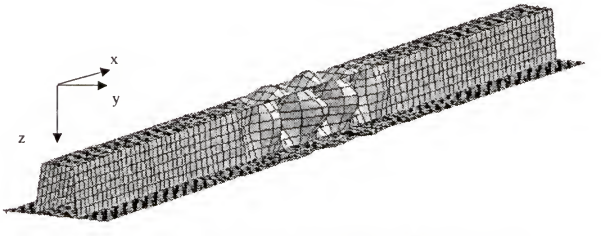
(a) Design Point 3, combined internal pressure and in-plane load case



(b) Design point 5, internal pressure load case



(c) Design point 11, combined internal pressure and in-plane compression load case



(d) Design point 2, combined load case

Figure 22: Examples of buckling modes predicted by STAGS analyses. Design point numbers refer to Table 9

The first buckling mode shape for the configuration defined by design point 3 and subjected to combined internal pressure and spanwise compression loading is shown in

Figure 22a. The buckling is mostly confined to the cap of the hat stiffener in the interior section of the panel. The first buckling mode for the configuration defined by design point 5, and subjected to internal pressure only is shown in Figure 22b. For this design and load case buckling is confined mostly to the skin near the panel ends. Figure 22c presents an example of skin buckling in the thin section of the skin near the skin thickness discontinuity. Interactive buckling modes were also predicted, but the buckling load factors for these modes were substantially greater than 1.0. An example of an interactive buckling mode shape is shown in Figure 22d, where both the cap and the webs of the hat stiffener buckle simultaneously.

Since individual components of the hat-stiffened panel buckle independently for buckling load factors close to 1.0, an approximate buckling analysis was developed by representing a subregion of each component of the hat stiffened panel as a simply supported plate. These subregions were assumed to have width, w, equal to the width of the component of the stiffener they were representing (e.g., the width $w = 4.3$ *in.* was used for the crown of the hat stiffener), and length equal to the length of a half wave of the bucking mode that corresponds to the lowest critical load value of a plate of width w that is simply supported on all edges. The component with the lowest buckling load determined the buckling load for the hat-stiffened panel.

The spanwise stresses in the panel induced by the internal pressure load vary quadratically along the length of the panel. However, the stress is approximately constant over a single buckling half wave, since the length of a single buckling half wave is much smaller than the length of the panel, (see Figure 22). Therefore, the buckling stress for a

single component of the hat stiffened panel was approximated from the buckling solution for a simply supported plate under constant load, that is given by:

$$\sigma_{buck} = \frac{\pi^2}{t\, m^2 b_b^{\,2} R^2}\left[m^4 D_{11} + 2\,(D_{12} + 2\,D_{66})(mnR)^2 + D_{22}(nR)^4\right] \tag{4.3}$$

where b_b and t are the width and the thickness respectively of the component of the hat stiffened panel under consideration, $R = \frac{a_b}{b_b}$ is the ratio between the length and the width of the panel component, and m and n are respectively the number of half waves in the direction parallel and perpendicular to the loading N_x. It can be shown that the minimum buckling stress for a simply supported plate with fixed width b_b and thickness t is achieved for a length a_b:

$$a_b = \left(\frac{D_{11}}{D_{22}}\right)^{1/4} b_b \tag{4.4}$$

Substituting Eq (4.4), and the material properties given in Table 1 into Eq. (4.3) gives the approximate buckling stress for an individual component of the hat stiffened panel:

$$\sigma_{buck} = 2.25\ 10^8 \left(\frac{t}{b_b}\right)^2 \tag{4.5}$$

An approximation for the buckling load factor was obtained by dividing the approximate buckling stress given by Eq. (4.5) by the approximate applied stress, determined from the simple beam analysis discussed above. The approximate buckling load factor is defined as $\lambda = \frac{\sigma_{buckling}}{\sigma_{applied}}$. A value of b_b of 8.0 inches was used in the calculations of σ_{buck} for the skin near the ends of the panel and for the skin at the thickness discontinuity ($b - w$ in Figure 21). To evaluate the buckling stress in the crown of the hat stiffener, the value of b_b was set equal to 4.3 inches (w in Figure 21).

Comparison of the buckling load factor obtained using the approximate analysis with the buckling load factor predicted by STAGS for the 26 designs provided in Table 9 indicated that the approximate buckling load factor underestimated the panel buckling resistance. This result was expected since the neighboring plate elements provide more boundary restraint than the one provided by simple support boundary condition. Also the stresses are not constant over the length of a component subregion, and the buckling load factor was calculated in the approximate analysis on the basis of the maximum stress over the length a_b. The buckling load factors predicted by the approximate analysis were up to 75% smaller than the buckling load factors predicted by STAGS. To improve the buckling load factor predictions obtained using the approximate beam-plate analysis, the STAGS results were used to fit a scale factor (as a single-term response surface approximation) to the buckling load factor in each critical region of the panel. The computed scale factors are provided in Table 10.

Table 10: Scale factors for combined load case buckling load factor λ_c

Buckling location	Scale factor
Cap in the interior section	1.804
Skin in the interior section	3.943
Skin at the end section	1.313

The combined load condition was critical for buckling for all of the designs considered, so the response surface was generated for this load case only. The buckling load factors predicted by the approximate analysis combined with the scale factors provided in Table 10, had a maximum error of 10% for panels whose buckling deformations were primarily localized to one component of the hat-stiffened panel. The approximate analysis generally overestimated the buckling load factor for panels that

activate an interactive buckling response. Fortunately these structures had buckling load factors that were well above 1.0.

Minimum Weight Structure.

The four three-level design variables, (d, t_{sb}, t_{cb}, t_{cm}) and two-level design variable t_w, generate a design domain of 162 points. All 162 designs were analyzed using beam theory to determine the stresses, and the scaled plate-beam approximation to determine the buckling load factor. The most promising design was identified and checked by STAGS analysis. The optimum design, found by inspection of the approximate analysis predictions, is given in Table 11.

Table 11: Minimum weight structural design

Design Variable	Value
d (in)	24
t_{sb} (in)	0.165
t_{cm} (in)	0.165
t_{cb} (in)	0.220
t_w (in)	0.165
Weight (lb/ft^2)	2.755

The structure shown in Table 11 weighs *2.75 lb/ft*^{2}. The maximum spanwise stress predicted by the approximate analysis was a tensile stress of 44,101 psi, in the crown of the hat stiffener, for the pressure-only load condition. The minimum approximate buckling load factor was for the combined load condition $\lambda = 0.97$. The buckling load corresponds to skin buckling in the interior region of the panel. The design provided in Table 11 was further checked for adequacy using a STAGS linear bifurcation buckling analysis from both a linear prestress state and a nonlinear prestress state. Both analyses were concluded for the combined load case, with both the internal pressure load and the

in-plane compression load applied in one load set. The buckling load factor obtained from the analysis is therefore applied to both loads. The STAGS linear bifurcation buckling analysis from a linear prestress state predicted a buckling load factor, $\lambda = 0.93$. The buckling mode was predominantly buckling of the skin at the axial location where the cross-section changes and is consistent with the buckling mode predicted by the approximate analysis. The predicted maximum tensile stress in the crown of the hat stiffener near the ends was 45,097 psi.

A linear bifurcation analysis executed from a nonlinear prestress state, and with both loads applied on one load set, predicted a buckling load factor $\lambda_c = 1.11$. The corresponding buckling mode shape is shown in Figure 23.

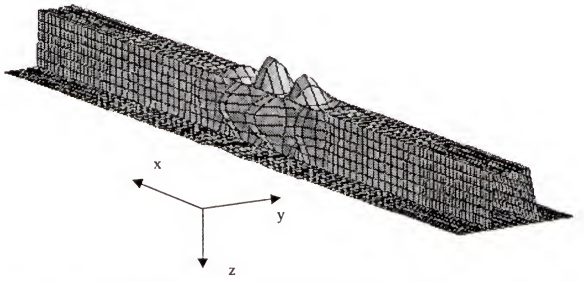


Figure 23: Linear bifurcation buckling mode from a nonlinear prestress state of optimal design for combined loading condition

The buckling mode is interactive, and the crown and the web of the hat stiffener buckled simultaneously. The interactive buckling mode predicted by the bifurcation analysis conducted from a nonlinear prestress state λ_c corresponds to the ninth buckling mode found by the bifurcation analysis conducted from a linear prestress state and its value was $\lambda_c = 1.21$. The buckling load factor predicted by the bifurcation analysis from a nonlinear prestress state is 14% higher then the buckling load factor predicted from a bifurcation buckling analysis executed from a linear prestress state.

Concluding Remarks for Structural Optimization of the Hat Stiffener Panel

A design study for structural optimization of a hat-stiffened laminated composite panel concept for the upper cover panel of a typical passenger bay of a blended-wing-body transport plane is presented. The hat-stiffened panel was subjected to internal pressure load, and to combined internal pressure and in-plane loads. The structural optimization problem was formulated using weight as the objective function with stress and buckling constraints. The initial optimization was conducted for a design having uniform properties in the spanwise direction using the PANDA2 program. The design was then refined using a more flexible design procedure that combined the STAGS finite element analysis program with response surface approximations for the stresses and for the buckling loads.

The design loads and safety factors were changed during the design process, but the knowledge of the design space gained from the first load case allowed an easy update to the optimum design. Furthermore, in the small region in the design space that was used for updating the design, simple beam and plate approximations were used with simple correction factors obtained from STAGS to generate response surface approximations for the stresses and for the buckling loads.

The design study demonstrated the use of three levels of analysis models and the use of response surface approximations for integrating the analysis models into the design process. This process resulted in a reduction in the weight of the hat-stiffened laminated composite panel concept from *4.30 lb/ft*^{2} to *2.75 lb/ft*^{2}. The hat stiffener panel concept proved to be a valid alternative to the thick sandwich concept since it achieved a weight of less than *3.0 lb/ft*^{2}.

CHAPTER 5
MULTIFIDELITY DESIGN OF A COMPOSITE PANEL WITH A CRACK

Summary

Crack propagation is an important concern in the design of aircraft composite fuselage and wing panels. However, computer simulation of crack propagation is computationally expensive. This chapter proposes combining high-fidelity models of analysis and low-fidelity models to calculate the crack propagation constraint in the design optimization process. The strategy uses correction response surfaces to relate the high-fidelity models to the low-fidelity models. Four different forms of correction response surface approximation are explored and their prediction capabilities compared. The multifidelity approach is found to be more accurate than single fidelity response surface method for the same computational cost.

Introduction

The next generation of supersonic transport planes are expected to fly at very high altitudes and are very vulnerable to pressure loss due to large cracks, such as due to turbine blade penetrations. Consequently, accurate prediction of crack propagation is a major issue in the structural design of such aircraft. Accurate calculation of crack propagation requires detailed structural models for computing stress intensity factors, which may be too expensive computationally for structural design optimization.

The problem of optimization with high cost models is common to many engineering design problems. Recently, there has been growing interest in taking

advantage of the simple models that were in use a generation ago when computers were less powerful. These models, which are less accurate, are termed here *low-fidelity* (LF) *models*. Combining the low-fidelity models with more accurate but expensive *high-fidelity* (HF) *models* can provide a good combination of high accuracy and low cost.

A common approach of such *multifidelity* techniques is to use the ratio or difference between the low-fidelity and high-fidelity models at one or several points in order to correct the low-fidelity model at other points. Recent examples of such multifidelity techniques are presented in *Toropov et al.* (1999). *Haftka* (1991) and *Chang et al.* (1993) suggested calculation of the ratio and its derivatives at one point in order to construct a linear approximation of the ratio at other points in design space. Fitting the ratio or the difference as a response surface approximation has the advantage over derivative based approximation that it can apply over larger regions in the design space and also smooth out numerical noise.

The *correction response surface* approach was used for combining high-fidelity Euler solutions with lower-fidelity panel solution for the design optimization of the high speed civil transport (HSCT) by *Knill et al.* (1999). *Mason et al.* (1994) in a different study on the HSCT instead combined lower-fidelity 2-D finite element (FE) models, and higher-fidelity 3-D FE. *Balabanov et al.* (1998) combined coarse and refined FE models to predict the optimum wing bending material weight of the HSCT as a function of the aircraft geometry. *Kaufman et al.* (1996) used a different strategy to combine low-fidelity and high-fidelity models for the optimization of the takeoff gross weight for the HCST. They used inexpensive constraints to eliminate from consideration large portions of the design box defined by the upper and lower limits on the design variables. Response

surface approximations were then constructed in the reduced design domain based on high-fidelity methods only. *Vitali et al.* (1997) combined strength-of-materials models with STAGS FE models to predict stresses and buckling loads in a hat-stiffened panel. *Vitali et al.* (1998) and *Vitali et al.* (1999) demonstrated the use of a correction response surface for calculating stress intensity factors for cracks in stiffened panels. The present work summarizes this earlier work and extends it in that response surfaces also approximate the low-fidelity model. The high-fidelity model describes the crack in the stiffened panel with a detailed finite element model near the crack and computes the stress intensity factor using classical linear elastic fracture mechanics. The low-fidelity model uses a coarse finite element model that does not model the crack. The stress intensity factor is calculated from a closed-form solution.

In one approach, a linear correction response surface is fitted to the values of the ratio and the difference of the high-fidelity and low-fidelity stress intensity factors and used with a quartic response surface based on the low-fidelity analyses. In a second approach, quartic response surface approximations are constructed based on the ratio and difference approximations for the corrected low-fidelity stress intensity factors in terms of the design variables. The quartic response surface approximation is used as a constraint in the design optimization problem.

A comparison is also carried out between the results obtained by using the correction response surface approach and a single-fidelity response surface approach.

Correction Response Surface

Correction response surfaces techniques couple high-fidelity and low-fidelity methods of analysis at some points in the design domain. In particular two ways of

coupling the two methods of analysis are considered in this paper: the ratio and the difference.

The process starts with high-fidelity and the low-fidelity computation of the response of interest, y, at n design points. The response y is a function of the vector of design variables $\mathbf{x}$ that is $y = y(\mathbf{x})$. At the n design points the ratio β

$$\beta = \frac{y_{HF}}{y_{LF}} \tag{5.1}$$

and the difference δ

$$\delta = y_{HF} - y_{LF} \tag{5.2}$$

are then computed. The subscript HF and LF in Eqs. (5.1) and (5.2) indicate the value of y obtained from a high-fidelity and a low-fidelity analyses, respectively. Using the n values of β and δ, two response surfaces approximations can be obtained as a function of the design variable vector $\mathbf{x}$, that is

$$\beta^{res} = \beta^{res}(\mathbf{x}) \text{ and } \delta^{res} = \delta^{res}(\mathbf{x}) \tag{5.3}$$

The multifidelity approximation to y at any other point is obtained from a low-fidelity analysis, y_{LF} as

$$y \cong \beta^{res}(\mathbf{x})\, y_{LF} \text{ or } y \cong \delta^{res}(\mathbf{x}) + y_{LF} \tag{5.4}$$

Implementing, the expressions of Eq. (5.3) and Eq. (5.4) for design optimization requires coupling of an optimizer with the low-fidelity analysis program. Every time the optimizer needs a value of y, it has to call the low-fidelity analysis program, and correct the low-fidelity solution with the correction response surface (β^{res} or δ^{res}).

In order to avoid the need to interface the low-fidelity analysis and the optimization software, in this chapter a different approach was investigated. The low-

fidelity analysis is replaced with a high order polynomial response surface approximation. The optimizer then calls the response surface approximation instead of calling the low-fidelity analysis program every time a value of y is needed. It is important to point out that the response surface approximation could be more accurate than the low-fidelity analysis itself as response surfaces tend to filter out the inherent numerical noise present in many computer simulations. Using the response surface of the low-fidelity analyses, $y_{LF}^{res}(x)$ and the ratio correction response surface approximation the value of y can then be approximated by

$$y(x) = \beta^{res}(x)\, y_{LF}^{res}(x) \tag{5.5}$$

or alternatively using the difference correction response surface approximation as

$$y(x) = \delta^{res}(x) + y_{LF}^{res}(x) \tag{5.6}$$

In addition to the approximations given by Eq. (5.5) and Eq. (5.6) two other approximate expressions of y can be obtained by correcting the low-fidelity values y_{LF} by β^{res} and δ^{res} and then fitting a high order polynomial response surface approximation to the corrected values of y_{LF}.

Example Problem Description

The dimensions of a composite stiffened panel used to demonstrate correction response surface techniques are shown in Figure 24. The composite stiffened panel was subjected to a tensile load case of 2000 lb/in.

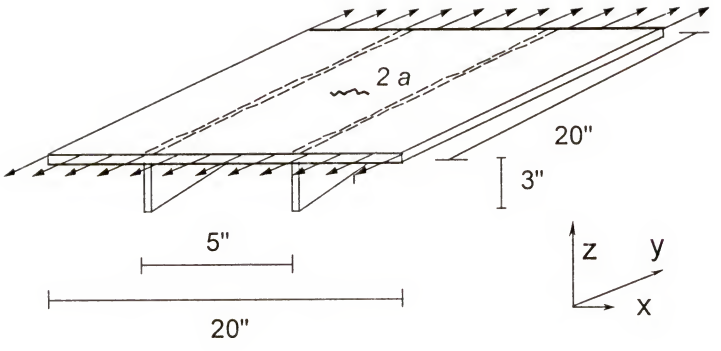


Figure 24: Wing panel geometry with a crack

The material used in the panel was AS4/3501-6 graphite/epoxy, and its properties are given in Table 12.

Table 12: Material properties of AS4/3501 -6

Material property	Value
Young's modulus E_1	20×10^6 psi
Young's modulus E_2	1.4×10^6 psi
Shear modulus G_{12}	0.76×10^6 psi
Poisson's Ratio ν_{12}	0.30
Density ρ	0.057 lb/in^3
K_q^0 (Fracture toughness)	100.000 psi$\sqrt{\text{in}}$

The geometry of the panel and the size of the crack ($2a$ = 4.0 inches), as well as the stacking sequence of the skin and the stacking sequence of the stiffener were kept constant throughout the study. The lay-up used for both the skin and the stiffener of the panel was $[\pm 45/90/0]_s$. The thicknesses of different plies, present in the skin and in the stiffener were used as design variables. The thicknesses of the +45° and –45° plies in each laminate were set equal to each other to maintain balance. A total of six ply thickness design variables were used and were labeled as: p_{45}, p_{90}, p_{00}, b_{45}, b_{90}, b_{00}. The last two digits indicate the ply angle while the letter "p" and the letter "b", respectively, refer to

the plate (skin) or the blade (stiffener). All design variable values ranged between 0.005 inches and 0.025 inches. The lower limit reflects that the material system used in this study, the AS/3501-6, is generally available in the form of prepreg tape with a nominal thickness of 0.005 inch. The maximum values represent estimates of the required material based on the papers by *Vaidya and Sun* (1996) and *Vaidya et al.* (1997).

Finite Element Models

The stiffened panel is simply supported along the edges parallel to the x–axis, and free along the edges parallel to the y–axis (see Figure 24). A multipoint constraint was applied simultaneously to the skin and to the stiffener, enforcing the same displacement in the y-direction along the simply supported edges. Two different finite element models of the stiffened panel were created using the STAGS program. In both finite element models, symmetry was utilized in order to model only ¼ of the structure, thereby reducing the total number of degrees of freedom.

The high-fidelity STAGS model of the stiffened panel was used to accurately capture the stress gradients in the region near the crack tip. The refined model consisted of 7,158 4-node elements and 46,680 degrees of freedom and is shown in Figure 25.

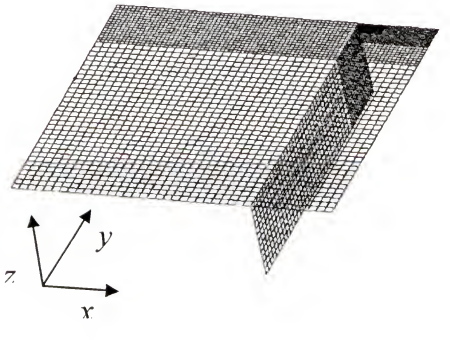


Figure 25: High-fidelity finite element model of ¼ of wing panel

The low-fidelity finite element model was coarser and it did not contain a crack. A uniform mesh of 4-noded elements for a total of 13284 degrees of freedom was used. The model was used to compute the nominal strains and stresses in the location where a crack is assumed to be. Then, a closed form solution was used to compute the approximate stress intensity factor in the presence of the crack.

Stress Intensity Factor Calculation

The method of *Vaidya and Sun* (1996) was used to predict crack propagation in a composite laminate. In this method the average stress intensity factor, K, through the thickness of the plate is computed by modeling the laminate as an equivalent orthotropic plate. Then the stress intensity factor in the 0^0 ply is computed using

$$K^0 = \eta K \tag{5.7}$$

where η is defined as the ratio of the nominal stress in the 0^0 ply to the average stress in laminate:

$$\eta = \frac{\sigma_{yy}^0}{\sigma_{yy}} \tag{5.8}$$

The crack in the skin of the panel is assumed to propagate, leading to structural failure, when stress intensity factor in the 0^0 plies K^0 reaches K_q^0, the fracture toughness of the 0^0 ply. The value of the fracture toughness K_q^0 was taken from *Vaidya and Sun* (1996). The theoretical distribution of σ_{yy} near the crack tip of the equivalent orthotropic panel is given by:

$$\sigma_{yy} = \frac{K}{\sqrt{2\pi r}} \tag{5.9}$$

The high-fidelity value of K, K_{HF}, was then obtained by least-square fit of σ_{yy} from FE data as a function of the distance from the crack tip, r using Eq. (5.9). Calculation of one K_{HF} required about 700 seconds of wall time and 51.7 seconds of CPU time on a 233 MHz DEC Alpha Station 200 4/233.

The low-fidelity stress intensity factor K_{LF} was calculated using the low-fidelity FE model and the infinite cracked plate solution for K

$$K = \sigma_{yy} \sqrt{\pi a} \tag{5.10}$$

where σ_{yy} in Eq. (5.10) is the average stress in an uncracked equivalent orthotropic panel at the location where the crack is assumed to be present. Calculation of one value of K_{LF} required an average of 11 seconds of wall time and 2.6 seconds of CPU time on a 233 MHz DEC Alpha Station 200 4/233. The low-fidelity value of K obtained is approximate mainly for two reasons: (i) the finite size of the panel was not taken into consideration in Eq. (5.10) and (ii) the beneficial effect of the stiffener on reducing the stress intensity factor could not be fully captured by this model.

Linear Correction Response Surfaces Approximations

The high-fidelity and low-fidelity methods of analysis were coupled together to model the crack propagation constraint. The motivation for coupling the two analysis methods was to try to achieve the accuracy of the high-fidelity analysis with the low computational cost of the low-fidelity analysis.

The original six design variables p_{45}, p_{90}, p_{00}, b_{45}, b_{90}, b_{00} were coded so that they ranged from –1 to +1. The coded design variables that correspond to the original design variables were named P_{45}, P_{90}, P_{00}, B_{45}, B_{90}, and B_{00}. A $2^{(6-2)}$ point fractional factorial design in the 6 coded design variables was selected using the JMP program for a total of

16 design points. At these 16 design points, the value of the stress intensity factor K^0 was calculated using both the high-fidelity method of analysis K^0_{HF} and the low-fidelity methods of analysis K^0_{LF}. Based on the values of K^0_{HF} and K^0_{LF} the ratio

$$\beta = \frac{K^0_{HF}}{K^0_{LF}} \tag{5.11}$$

and the difference

$$\delta = K^0_{HF} - K^0_{LF} \tag{5.12}$$

was computed at all 16 design points. The 16 values of the ratio β varied between 0.86 and 1.01 indicating moderate differences between the high-fidelity and the low-fidelity model. The 16 values of the difference δ varied between 17.06 psi$\sqrt{\text{in}}$ and 16811 psi$\sqrt{\text{in}}$ for bladed stiffened panels having a high-fidelity stress intensity factor K^0_{HF} of 67106 psi$\sqrt{\text{in}}$ and 109510 psi$\sqrt{\text{in}}$, respectively. The largest discrepancies for the ratio β and the difference δ among the 16 design points did not occur at the same design points, however they both occurred at design points that corresponded to bladed stiffened panels with a very thin skin and a very thick stiffener.

One linear response surface approximation for the ratio and one linear response surface approximation for the difference, β^{Lin} and δ^{Lin} respectively, were obtained fitting the values 16 values of β and δ as a function of P_{45}, P_{90}, P_{00}, B_{45}, B_{90}, and B_{00}. The expressions for β^{Lin} and δ^{Lin} found from a least square fit on the 16 values of β and δ were

$$\begin{aligned}\beta^{Lin} = 0.01\ (93.7 + 2.39\ P_{45} + 0.612\ P_{90} + 1.48\ P_{00} + \\ - 1.63\ P_{45} - 0.479\ P_{90} - 2.21\ P_{00})\end{aligned} \tag{5.13}$$

and

$$\delta^{Lin} = -7009 + 3620\,P_{45} + 1108\,P_{90} + 3856\,P_{00} + \\ -1455\,P_{45} - 337\,P_{90} - 1344\,P_{00} \tag{5.14}$$

Once the expressions for the correction response surface approximation were obtained, the stress intensity factor in the 0^0 plies for any other design can estimated by.

$$K^0 \cong \beta^{Lin}\,K^0_{LF} \tag{5.15}$$

or by

$$K^0 \cong \delta^{Lin} + K^0_{LF} \tag{5.16}$$

An additional linear response surface approximation based only on the high-fidelity stress intensity factors of the previous 16 design points was constructed and named $K^{0,Lin}_{HF}$

$$K^{0,Lin}_{HF} = 10^4\,(9.68 \quad -1.67\;\; P_{45} - 0.586\;\; P_{90} - 3.53\;\; P_{00} + \\ -1.19\;\; P_{45} - 0.336\;\; P_{90} - 1.80\;\; P_{00}) \tag{5.17}$$

The accuracy of the fit was checked by the statistical measures presented in Chapter 3 at 48 additional high-fidelity analyzes at the points that complement the design to a full two level factorial design.

The important statistical parameters of β^{Lin}, δ^{Lin} and $K^{0,Lin}_{HF}$ are presented in Table 13. The $\%RMSE_{\mathrm{OR}}$ in Table 13 is defined by Eq. (3.31) with $\chi = 9$ while $\%RMSE_{\mathrm{CH}}$ is again defined by Eq. (3.31) with $\chi = 48$. Table 13 also reports the average values of the stress intensity factor found for each different approximations and the average error of each different approximation. As can be seen from Table 13 the most accurate approximation is given by β^{Lin} and the least accurate one is given by $K^{0,Lin}_{HF}$. The statistical parameters of $K^{0,Lin}_{HF}$ indicate that the response surface is not accurate. This is

confirmed by the optimization problem results presented later where $K_{HF}^{0,Lin}$ was used as crack propagation constraint.

Table 13: Accuracy measures of the correction response surfaces β^{Lin}, δ^{Lin} response surface $K_{HF}^{0,Lin}$ obtained using only high-fidelity analyses The subscript OR and CH denote the original set of 16 points and the independent set of 48 check points, respectively

Parameter	β^{Lin}	δ^{Lin}	$K_{HF}^{0,Lin}$
R^2	0.954	0.897	0.814
R^2_{Adj}	0.923	0.830	0.701
RMSE	$1.16\ 10^{-3}$	2420	$2.82\ 10^{3}$
$\%RMSE_{OR}$ (Eq. 8) χ=16	1.240	35.10	29.10
$\%RMSE_{CH}$ (Eq. 8) χ=48	0.858	37.10	18.50
Av. $(K^0_{Approximate})_{OR}$	96729	96842	96843
Av. $(Error)_{OR}$ of K^0	734	1503	17514
Av. $(K^0_{Approximate})_{CH}$ K^0	95524	95662	95107
Av. $(Error)_{CH}$	715	2150	14460
Smallest t_{stat}	1.670	0.750	0.520
p-value	0.129	0.460	0.616

The high p-values for the smallest t_{stat} of both δ^{Lin} and K_{HF}^{Lin} indicate that they have a term, which turns out to be B_{90}, that has a high probability of being equal to zero. Removing the B_{90} term from the expressions of δ^{Lin} and $K_{HF}^{0,Lin}$ however did not increase the quality of the approximations and therefore the term was kept in the response surface expressions of δ^{Lin} and $K_{HF}^{0,Lin}$.

Response surface approximation based on the low-fidelity model

Combining the correction response surfaces β^{Lin} and δ^{Lin} and the values of the stress intensity factor obtained from the low-fidelity model K_{LF}^{0} provides an accurate approximation to the crack propagation constraint. However, this combination still

requires coupling of the optimizer with the low-fidelity analysis program as explained in the section "Correction response surface". In order to eliminate this requirement, the low-fidelity analyses were approximated by with a fourth order polynomial response surface..

The 400 design points were chosen using the latin hypercube (LHS) design in the six coded design variables P_{45}, P_{90}, P_{00}, B_{45}, B_{90}, and B_{00} assuming a uniform distribution for each design variable. LHS is a space filling design and it tends to space design points evenly, letting each design variable take a particular value only once. However, response surfaces based on LHS designs are often not very accurate at corners of the design space. Therefore a full two-level factorial design was added to the LHS design for a total of 464 design points. The K_{LF}^{0} was computed at the 464 points and a quartic response surface in the six coded design variables was obtained. After stepwise regression was carried out 148 out of the original 210 terms were retained for a C_p = 114. The statistical parameters of the resulting quartic response surface $K_{LF}^{0,Quartic}$ for the low-fidelity model are presented in Table 14.

Table 14: Statistical properties of quartic response surface approximation of $K_{LF}^{0,Quartic}$ obtained using 464 low-fidelity analyses. The subscript OR and CH denote the original set of 464 points and the independent set of 4032 check points, respectively

Parameter	$K_{LF}^{0,Quartic}$
N of Terms	148
C_p	114
R^2	0.999
R^2_{Adj}	0.999
RMSE	365
$\%RMSE_{OR}$ χ=462	0.404
$\%RSME_{CH}$ χ=4032	1.430
Smallest t_{stat}	1.330
p value	0.184

In addition Figure 26 shows the prediction of the quartic response surface approximation based on 464 low-fidelity analyses compared with the values from 4032 additional low-fidelity analysis. The 4032 additional designs were obtained by subtracting the 64 designs that constitute a two-level factorial design from the 4096 designs that complete a four-level factorial design. If the quartic response surface approximation were perfectly accurate the values represented by dots in Figure 26 would lie on the main diagonal. The figure indicates that the response surface approximation is very accurate, allowing us to use with confidence the quartic response surface approximation $K_{LF}^{0,Quartic}$ instead of the low-fidelity analysis in the optimization procedure.

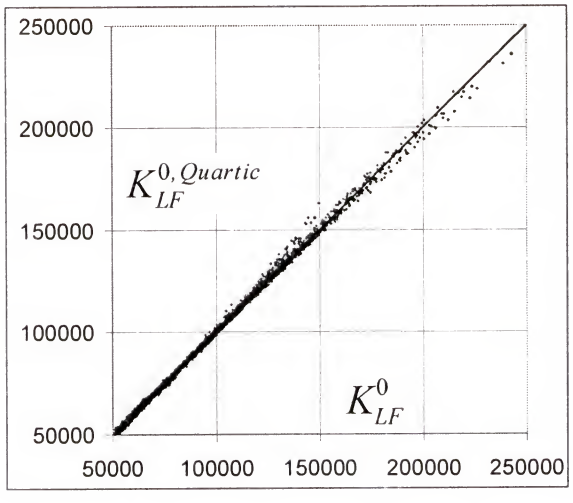


Figure 26: Quartic response surface approximation of the stress intensity factor, $K_{LF}^{0,Quartic}$ based on 464 low-fidelity analyses versus the stress intensity factor obtained from low-fidelity analysis K_{LF}

Stress intensity constraint approximation

Two different strategies were used to couple the correction response surfaces β^{Lin} and δ^{Lin} with the 464 low-fidelity analyses in order to obtain the expression of the stress

intensity factor to use in the optimization procedure, resulting in a total of four different approximations for the crack propagation constraint.

The first expression for the stress intensity factor approximation, $K^{0,\beta}$, was obtained by multiplying $K_{LF}^{0,Quartic}$ obtained in the previous section by β^{Lin}, that is

$$K^{0,\beta} = \beta^{Lin}\ K_{LF}^{0,Quartic} \tag{5.19}$$

while the second expression for the stress intensity factor constraint, $K^{0,\delta}$, was obtained by adding $K_{LF}^{0,Quartic}$ to δ^{Lin} as

$$K^{0,\delta} = \delta^{Lin} + K_{LF}^{0,Quartic} \tag{5.19}$$

In addition two other expressions of the stress intensity factor approximation were obtained by correcting the 464 K_{LF}^{0} values used in the previous section to fit $K_{LF}^{0,Quartic}$ using β^{Lin} and δ^{Lin}, and then fitting a quartic response surface to the 464 corrected values of the stress intensity factor. The response obtained by multiplying the K_{LF}^{0} values by β^{Lin} was named $K_{Quartic}^{0,\beta}$ and for the response obtained from adding K_{LF}^{0} to δ^{Lin} was named $K_{Quartic}^{0,\delta}$. The statistical parameters of the quartic response surfaces obtained, $K_{Quartic}^{0,\beta}$ and $K_{Quartic}^{0,\delta}$, are presented in Table 15.

Table 15: Quartic response surface approximations statistical parameters of $K^{0,\beta}_{Quartic}$ and $K^{0,\delta}_{Quartic}$ obtained from correcting K_{LF}, with β^{Lin} and δ^{Lin}. The subscript OR denotes the original set of 16 points.

Parameter	$K^{0,\beta}_{Quartic}$	$K^{0,\delta}_{Quartic}$
N of Terms	151	148
C_p	116	114
R^2	0.999	0.999
R^2_{Adj}	0.999	0.999
RMSE	341	365
$\%RMSE_{OR}$	0.400	0.440
Smallest t_stat	1.000	0.090
p value	0.317	0.929

In order to assess in a more reliable way the accuracy of the four expressions of the stress intensity factor approximations $K^{0,\beta}$, $K^{0,\delta}$, $K^{0,\beta}_{Quartic}$ and $K^{0,\delta}_{Quartic}$ the prediction of the approximations were compared with the stress intensity factor obtained from the high-fidelity method of analysis for the original 16 design points used to construct the correction response surfaces β^{Lin} and δ^{Lin} and for the additional 48 points. The results of the comparison are presented in Table 16.

Table 16: Comparison of the accuracy of the stress intensity factor constraints obtained by different methods. $K^{0,\beta}_{Quartic}$ is the quartic approximation obtained from 464 β^{Lin} corrected values. $K^{0,\delta}_{Quartic}$ is the quartic approximation obtained from 464 δ^{Lin} corrected values. $K^{0,\beta}$ is the approximation obtained by multiplying the response surface β^{Lin} by the quartic approximation based on 464 low-fidelity analyses. $K^{0,\delta}$ is the approximation obtained by adding the response surface β^{Lin} to the quartic approximation based on 464 low-fidelity analyses. The subscript OR and CH denote the original set of 16 points and the independent set of 48 check points, respectively.

Parameter	$K^{0,\beta}_{Quartic}$	$K^{0,\delta}_{Quartic}$	$K^{0,\beta}$	$K^{0,\delta}$
$\%RMSE_{OR}$ χ=16	1.68	3.67	1.68	2.82
$\%RMSE_{CH}$ χ=48	1.11	3.29	1.11	2.67
$\%Err\_Max_{OR}$ χ=16	1.39	6.08	1.41	6.58
$\%Err\_Max_{CH}$ χ=48	1.98	5.31	1.96	4.81

As can be seen from Table 16 the four approximations $K^{0,\beta}$, $K^{0,\delta}$, $K^{0,\beta}_{Quartic}$ and $K^{0,\delta}_{Quartic}$ differ in their accuracy. In particular it appears that the most accurate approximation is given by the approximation based on β^{Lin}, $K^{0,\beta}$ However from the results shown in the first four columns of Table 16, it appears that there is no significant difference in the prediction of the stress intensity factor approximations if the correction response surfaces are multiplied ($K^{0,\beta}$) or added ($K^{0,\delta}$) to the quartic response surface based on low-fidelity analyzes or if the low-fidelity data are first corrected and then a quartic response surface is fitted to the corrected data ($K^{0,\beta}_{Quartic}$ and $K^{0,\delta}_{Quartic}$).

Response surface approximation based only on high-fidelity method

Coupling the high-fidelity method and the low-fidelity method of analysis had a computational cost. In particular the total amount of CPU time required by 16 high-fidelity analysis and 464 low-fidelity analysis was about 2080 sec. The same amount of CPU time is required to perform about 40 high-fidelity analyses (if wall time is considered, than only 24 high-fidelity analyses can be performed) consequently 40 high-fidelity analyses were carried out for comparison. Sixteen high-fidelity analysis were already available from the construction of the correction response surfaces β^{Lin} and δ^{Lin}. Thirteen additional points were selected by choosing the design with all coded variable set to zero and then letting each design variable to take the values +1 and –1. Finally, 11 D-optimal points were chosen from the 48 points that with the original 16 complete a two level factorial design.

A quadratic response surface in the six coded design variables P_{45}, P_{90}, P_{00}, B_{45}, B_{90}, and B_{00}, called $K_{HF}^{Quadratic}$, was fitted to the high-fidelity results. Table 17 presents the most significant statistical parameters of $K_{HF}^{Quadratic}$, after the stepwise regression procedure based on C_p was applied.

Table 17: Response surface approximation $K_{HF}^{Quadratic}$ obtained from only high-fidelity analyses. The subscript OR and CH denote the original set of 40 points and the independent set of 37 check points, respectively

Parameter	$K_{HF}^{Quadratic}$
N of Terms	19
C_p	11.76
R^2	0.990
R^2_{Adj}	0.980
RMSE	6137
Smallest t_stat	1.420
p value	0.170
$\%RMSE_{OR}$ χ=40	6.460
$\%RMSE_{CH}$ χ=37	12.97
$\%Err\_Max_{OR}$ χ=40	9.070
$\%Err\_Max_{CH}$ χ=37	40.77

The prediction capabilities of $K_{HF}^{Quadratic}$ can compared with the other forms of approximations $K^{0,\beta}$, $K^{0,\delta}$, $K_{Quartic}^{0,\beta}$ and $K_{Quartic}^{0,\delta}$ shown in Table 16. In Table 18 the values of $K_{HF}^{Quadratic}$ labeled by "OR" refer to the 40 points used to construct the quadratic response surface approximation while the quantities labeled "CH" refer to the 37 points that with 29 out of the 40 original points complete the two level factorial design. Comparing the results presented in Table 16 and Table 18 it is clear that the response surface approximation $K_{HF}^{Quadratic}$ is not as accurate as the approximations obtained using the correction response surface approach.

Optimization

The minimum weight optimization formulation for the composite stiffened panel shown in Figure 25 is summarized in Table 18.

Table 18: Minimum weight optimization problem for the bladed stiffened panel

Objective Function:	Min{Weight}
Design Variables :	P_{45}, P_{90}, P_{00}, B_{45}, B_{90}, B_{00}
Constraints:	$-1 < P_{45} < +1$
	$-1 < P_{90} < +1$
	$-1 < P_{00} < +1$
	$-1 < B_{45} < +1$
	$-1 < B_{90} < +1$
	$-1 < B_{00} < +1$
	$K^0 < 100{,}000 \text{ psi}\sqrt{\text{in}}$

The weight optimization problem was implemented into the Excel (*Microsoft Corporation,* 1998) solver tool and repeated for each of the six approximations of the stress intensity factor $K^{0,\beta}$, $K^{0,\delta}$, $K^{0,\beta}_{Quartic}$, $K^{0,\delta}_{Quartic}$, K^{Lin}_{HF} and $K^{Quadratic}_{HF}$. The minimum weights obtained are summarized in Table 19.

Table 19: Optimum designs obtained from the different stress intensity factor approximations. %Err denotes the percentage error of each approximation at its optimum (negative when conservative)

Parameter	$K^{0,\beta}_{Quartic}$	$K^{0,\delta}_{Quartic}$	$K^{0,\beta}$	$K^{0,\delta}$	$K^{Quadratic}_{HF}$	K^{Lin}_{HF}
P_{45}	-1	-1	-1	-1	-1	-1
P_{90}	-1	-1	-1	-1	-1	-1
P_{00}	-0.290	-0.333-	-0.326	-0.355	-0.194	0.481
B_{45}	-1	-1	-1	-1	-1	-1
B_{90}	-1	-1	-1	-1	-1	-1
B_{00}	1	1	1	1	1	1
%Err	1.890	0.260	0.300	-0.880	6.270	34.5
Weight (lb)	0.446	0.441	0.442	0.438	0.457	0.533

Moreover the optimizer always found the minimum weight structure by setting all the non zero degree plies to the minimum thickness gage and then increasing the thickness of the 0^0 plies in the blade as much as possible. Once the upper bound on B_{00} was reached the optimizer started increasing the 0^0 plies in the panel, P_{00}.

The most accurate values at the optimum are obtained for the approximations based on δ^{Lin}. However all four approximations, $K^{0,\beta}$, $K^{0,\delta}$, $K^{0,\beta}_{Quartic}$ and $K^{0,\delta}_{Quartic}$ are very accurate and are in line with the results presented in Table 16. In addition it can be also noticed how the approximations based only on a single-fidelity methods $K^{Quadratic}_{HF}$ and $K^{Quadratic}_{HF}$ are less accurate than the multiple-fidelity methods.

Concluding Remarks

In many engineering design problems optimization presents the obstacle of requiring computationally expensive analyses. In order to overcome this difficulty more attention has been recently given to simple models that were used as design tools a generation ago. These models are often referred to as low-fidelity models and though not very accurate in nature they are computationally inexpensive. This study demonstrates

that combining the low-fidelity models with more accurate but expensive high-fidelity models it is possible to achieve high accuracy at a low computational cost.

A crack propagation constraint for composite stiffened panel designs is used as an example to demonstrate the feasibility of combining low-fidelity models and high-fidelity models to obtain accurate results at a low computational cost. Two linear polynomial correction response surface approximations were fitted to the ratio and the difference of the stress intensity factors obtained from a high-fidelity model and a low-fidelity model at a relatively small set of design points. Two different approaches were then used to combine the low-fidelity and high-fidelity methods. The first approach used the correction response surfaces together with a quartic response surface that approximated the low-fidelity analyses. The second approach instead corrected the values of the low-fidelity analyses and then fit a quartic response surface to the corrected values. The results obtained using the correction response surface approach proved to be more accurate than the results obtained using only high-fidelity analyses requiring the same amount of computational resources.

In addition there was not much difference if the ratio or the difference was used to approximate the crack propagation constraint. Moreover it did not matter if the low-fidelity analyzes were fitted with a quartic response surface and then coupled with the correction response surface or if the low-fidelity analyses were first corrected and then fitted with a quartic response surface approximations.

CHAPTER 6
RESPONSE SURFACE APPROXIMATIONS INCLUDING GRADIENT INFORMATION

Introduction

In the past decade, there has been growing interest in using response surface approximations with data generated by computer simulation, partly due to the realization that these simulations often have substantial numerical noise that can be reduced by the noise filtering properties of response surface approximations. In computer simulations, derivatives are often available and frequently at a lower computational cost. As a consequence, there is an interest in considering response surface approximations that incorporate both function and derivative data. Unfortunately it is not possible to apply the same methods used for the traditional response surface methodology since function and gradient values are strongly correlated. As a result, any methodology that attempts to use function and gradient information should take into consideration the strong coupling existing between the two types of data.

A similar situation of strongly correlated forms of data is encountered in the analysis of multiresponse experiments. An experiment in which a number of responses are measured simultaneously, for each setting of a group of input variables, is called a multiresponse experiment. The analysis of data from a multiresponse experiment requires careful consideration of the multivariate nature of the data, and as a consequence, the responses should not be investigated individually and independently of each other. Interrelationships that may exist among responses can render such univariate

investigation inaccurate. Several researchers including *Box et al.* (1970), *Khuri and Cornell* (1996), and more recently *Robinson* (2000) investigated the problems associated with creating a number of responses simultaneously. However multiresponse methodology cannot be applied directly to approximations based on function and gradients values. When using function and gradients of a quantity at the same time, problems may arise from the fact that not only there are several functions to be fitted at the same but additionally the fitting parameters are the same. In other words the design points and the parameters of each approximation (function or gradient) are the same while the shape functions that define the function and the gradient approximation are different.

Etman (1992), and subsequently *Van Keulen et al.* (2000), used function and gradient information to obtain response surface approximations. However in thier studies both *Etaman* and *Van Keulen et al.* had little or no control on the noise level in the data, and thus did not know the form of the true function. In this chapter an attempt is made to show the effectiveness of inclusion of gradient information in creating response surface approximations. Simple numerical problems were generated maintaining good control on the coupling between function and gradient values, noise level, and the form of the true function that it is approximated.

Regression analysis based on function and gradient values

Assume that y denotes the function that we want to approximate, and that it is a function of the design variables $\mathbf{x}$ with $\mathbf{x} \in \Re^{n_{dv}}$ where n_{dv} indicates the number of design variables, thus $y = y(\mathbf{x})$. The derivative of y corresponding to the direction $\mathbf{d}$, where $\|\mathbf{d}\|^2 = 1$, is denoted by y_d and is determined by

$$y_d = (grad^T y)\,\mathbf{d} \tag{6.1}$$

Consider n design points $\mathbf{x}_i$ (i = 1..n) where for each design point the function value $y(\mathbf{x}_i)$ and the γ directional derivatives exist. The associated derivative directions are denoted by d_{ij} (i=1..n, j=1..n_{dv}) and a compact notation can be introduced so that function values and the derivatives can be denoted by y_{ij}. The components y_{ij} are defined as

$$\begin{aligned} &y_{i0} \equiv y_i \equiv y(\mathbf{x}_i), && i = 1..n \\ &y_{ij} \equiv [grad^T y]_{x=x_i}\, d_{ij}, && j = 1..n \quad \text{and} \quad j = 1..\gamma \end{aligned} \tag{6.2}$$

Obviously if γ = 0 then the common case of response surface approximation based on only function values is obtained. The previous notation can be compacted even further by denoting

$$\mathbf{y}_i^T = [y_{i0}..y_{i\gamma}] \tag{6.3}$$

In principle, the length of y_i may vary from point to point, as γ is not necessarily constant for all points. In this work however the additional restrictions of $\gamma = n_{dv}$ and $\boldsymbol{d}_{ij} = \mathbf{e}_j$ with $\mathbf{e}_j$ being the basis vector in coordinate directions, are imposed.

The response function y will be approximated using an approximation $\hat{y}$ defined by

$$\hat{y} = \mathbf{f}(\mathbf{x})^T \boldsymbol{\beta} \tag{6.4}$$

where β is the vector of unknown parameters and $\mathbf{f}(\mathbf{x})$ is the vector of shape functions, usually monomials of polynomials. Directional derivatives follow as

$$\hat{y}_{\mathbf{e}_j} = \mathbf{g}^T \boldsymbol{\beta} \tag{6.5}$$

with

$$\mathbf{g} = \mathbf{g}(\mathbf{x}; \mathbf{e}_j) = (grad^T \mathbf{f}(\mathbf{x}))\,\mathbf{e}_j \tag{6.6}$$

Evaluation of the response surface approximation of Eq. (6.4) and its derivatives of Eq. (6.5) for $\mathbf{x}_i$, and the associated base vector $\mathbf{e}_j$ gives

$$\hat{y} = \mathbf{A}_i \boldsymbol{\beta} \tag{6.7}$$

where $\hat{y}$ is defined in an analogous way to Eq. (6.3). The matrix $\mathbf{A}_i$ is defined as

$$\mathbf{A}_i^T = [\mathbf{f}_i, \mathbf{g}_{i1}, \mathbf{g}_{i2}, .., \mathbf{g}_{i\gamma}] \tag{6.8}$$

with $\mathbf{f}_i = \mathbf{f}(\mathbf{x}_i)$ and $\mathbf{g}_{ij} = g(\mathbf{x}_i;\mathbf{e}_j)$.

The response surface of the function y_{i0} and its derivatives y_{ij} is

$$y_{ij} = \hat{y}_{ij} + \varepsilon_{ij} \tag{6.9}$$

where ε_{i0} refers to the error in the function value, and ε_{ij} for $j \neq 0$ refers to the errors in the directional derivatives. From direct generalization of the notation used, Eq. (6.9) can be written in the familiar form

$$\mathbf{y}_i = \hat{\mathbf{y}}_i + \boldsymbol{\varepsilon}_i \tag{6.10}$$

with ε_i being defined in an analogous way to Eq. (6.3).

The vector of regression coefficients $\boldsymbol{\beta}$ for the response surface approximation of Eq. (6.4) is determined using a weighted least square fit. Unlike standard weighted least square (WLS) methods in this study, information on both function values and derivatives are included. The weighting is introduced into the least square procedure by defining the inner product as

$$\mathbf{a} \bullet \mathbf{b} = \mathbf{a}^T \mathbf{W}\, \mathbf{b} \tag{6.11}$$

where in Eq. (6.11) **a** and **b** are two vectors and **W** is the weighting matrix. Thus, the least squares formulation is governed by

$$\psi = \sum_{i=1}^{n} (\mathbf{y}_i - \hat{\mathbf{y}}_i)^T \mathbf{W}_i (\mathbf{y}_i - \hat{\mathbf{y}}_i) \tag{6.12}$$

The estimates **b** for the regression coefficients $\boldsymbol{\beta}$ are selected as to minimize ψ. Substitution of Eq. (6.7) into Eq. (6.12) and minimization of the latter gives

$$\mathbf{b} = \left\{ \sum_{i=1}^{n} \mathbf{A}_i^T \mathbf{W}_i \mathbf{A}_i \right\}^{-1} \sum_{k=1}^{n} \mathbf{A}_k^T \mathbf{W}_k \mathbf{y}_k \tag{6.13}$$

In order to achieve a more compact notation, the following notation can be introduced

$$\mathbf{X}^T = [\mathbf{A}_1^T, ..., \mathbf{A}_n^T] \tag{6.14}$$

$$\mathbf{y}^T = [\mathbf{y}_1^T, ..., \mathbf{y}_n^T] \tag{6.15}$$

$$\mathbf{W} = \begin{pmatrix} \mathbf{W}_1 & 0 & \cdots & 0 \\ 0 & \mathbf{W}_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{W}_n \end{pmatrix} \tag{6.16}$$

and

$$\mathbf{H} = \{\mathbf{X}^T \mathbf{W} \mathbf{X}\}^{-1} \tag{6.17}$$

With these definitions Eq. (6.13) can be written in a compact form as

$$\mathbf{b} = \mathbf{H}\mathbf{X}^{\mathbf{T}}\mathbf{W}\mathbf{y} \tag{6.18}$$

Similarly to **y**, $\hat{\mathbf{y}}$ and $\boldsymbol{\varepsilon}$ can be defined as in Eq. (6.15). With these definitions it is possible to write an equation that has the familiar look of Eq. (3.21)

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \tag{6.19}$$

It should be noted that Eq. (6.19) also includes directional derivatives and therefore is fundamentally different from Eq. (3.21). In principle, the previous formulation can be generalized further by dropping the special form of the matrix **W** in Eq. (6.16) and by considering a fully populated matrix.

Statistical properties of gradient based response surfaces

When dealing with response surface approximation based only on function values, it is customary to assume that the data satisfy the Gauss-Markov conditions or

$$E(\boldsymbol{\varepsilon}) = 0 \tag{6.20}$$

$$\text{cov}(\boldsymbol{\varepsilon}) = \sigma^2 \, \mathbf{I} \tag{6.21}$$

in other words that the random errors have zero mean, are uncorrelated, and have the same variance σ^2. Moreover, it is also typically assumed that errors are distributed normally. Under these conditions, the least squares formulation provides the best linear unbiased estimate (BLUE) as in *Sen and Srivastava* (1990). When gradient information is included, the Gauss-Markov condition of Eq. (6.21) must be relaxed as

$$\text{cov}(\boldsymbol{\varepsilon}) = \overline{\mathbf{C}} \tag{6.22}$$

The matrix $\overline{\mathbf{C}}$ is assumed to be symmetric and positive definite. It is further assumed that the errors corresponding to different points are independent, that is $\text{E}(\boldsymbol{\varepsilon}_i \; \boldsymbol{\varepsilon}_j^T) = 0$ for $i \neq j$ so that $\overline{\mathbf{C}}$ reduces to a form similar to Eq. (6.16) that is

$$\overline{\mathbf{C}} = \begin{pmatrix} \mathbf{C}_1 & 0 & \cdots & 0 \\ 0 & \mathbf{C}_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{C}_n \end{pmatrix} \tag{6.23}$$

with $\mathbf{C}_i = \text{cov}(\boldsymbol{\varepsilon}_i \; \boldsymbol{\varepsilon}_i^T) = \text{E}(\boldsymbol{\varepsilon}_i \; \boldsymbol{\varepsilon}_i^T)$. The assumption on the normal distribution on $\boldsymbol{\varepsilon}$ is still retained. As $\text{E}(\boldsymbol{\varepsilon}) = 0$, it follows that

$$E(\mathbf{y}) = \mathbf{X}\boldsymbol{\beta} \tag{6.24}$$

and

$$\text{cov}(\mathbf{y}) = \overline{\mathbf{C}} \tag{6.25}$$

Under these assumptions it can be demonstrated that

$$E(\mathbf{b}) = \boldsymbol{\beta} \tag{6.26}$$

and the covariance of the parameters **b** can be calculated as

$$\text{cov}(\mathbf{b}) = \mathbf{H}\mathbf{X}^{\mathrm{T}}\mathbf{W}\overline{\mathbf{C}}\mathbf{W}^{\mathrm{T}}\mathbf{X}\mathbf{H}^{\mathrm{T}} \tag{6.27}$$

From Eq. (6.27) it follows that if the weighting matrix **W** is chosen as

$$\mathbf{W} = \overline{\mathbf{C}}^{-1} \tag{6.28}$$

the covariance of the parameters **b** is greatly simplified as

$$\text{cov}(\mathbf{b}) = (\mathbf{X}^{\mathrm{T}}\overline{\mathbf{C}}^{-1}\mathbf{X})^{-1} \tag{6.29}$$

It can be demonstrated that the vector of parameters **b** obtained using the relationship of Eq. (6.28) is again BLUE (*Sen and Srivastava,* 1990).

From *cov*(**b**) it is possible to calculate an estimate of the reliability of the **b** coefficients by taking the ratio between the value of a coefficient b_i and the square root of its variance *var*(b_i

$$t_{stat,i} = \frac{b_i}{\sqrt{\text{var}(b_i)}} \tag{6.30}$$

The t_{stat} can be used to reduce the number of parameters of a response surface approximation by eliminating the coefficient with the lowest t_{stat}. The procedure is repeated until all remaining coefficients have a t_{stat} that is higher than a specified limit. The coefficients are eliminated one at the time because when one coefficient is deleted the t_{stat} of all the other coefficients change. The limit can depend on the individual problem but good results were obtained in this chapter by using two as the limit.

Iterative procedure

A further simplification in the structure of the covariance $\overline{\mathbf{C}}$ can be made by assuming that the errors at different design points are uncorrelated but that the correlation between function and gradients at each design point is the same. That implies that the matrix $\overline{\mathbf{C}}$ of Eq. (6.23) can be written in the form

$$\overline{\mathbf{C}} = \mathbf{I}_n \otimes \mathbf{C} \tag{6.31}$$

where n is the number of design points and **C** is the function value-gradient covariance matrix. The subcovariance **C** is a symmetric $(1+n_{dv})$ x $(1+ n_{dv})$ matrix where n_{dv} is the number of design variables. Thus, $\overline{\mathbf{C}}$ is a n_t x n_t matrix where n_t is the total number of observations available (functions + gradients), that is $n_t = n\ (n_{dv}+1)$.

However, the covariance matrix $\overline{\mathbf{C}}$ is rarely known and therefore an estimate has to be used to achieve a proper weighting to minimize the *cov*(**b**) as in Eq. (6.29). The process of finding an estimate of $\overline{\mathbf{C}}$ leads to an iterative procedure since the covariance matrix $\overline{\mathbf{C}}$ depends on the vector of parameters **b**. However, **b** depends on **C** through the weighting matrix $\mathbf{W} = \mathbf{C}^{-1}$ used in the WLS procedure. Substituting Eq. (6.17) into Eq (6.18) the coefficients **b** are obtained as

$$\mathbf{b} = (\mathbf{X}^T \mathbf{W} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{W} \mathbf{y} \tag{6.32}$$

where **X** is the model matrix of Eq. (6.14), **W** is the weight matrix, and **y** is the vector of measurement as in Eq. (6.19). If there is no a priori information available about the covariance matrix, the identity matrix can be used as the starting weight matrix of an iterative procedure that attempts to estimate **C**. After solving the least squares problem of Eq. (6.32), **b** is used to calculate $\hat{\mathbf{y}}$:

$$\hat{\mathbf{y}} = \mathbf{X}\mathbf{b} \tag{6.33}$$

and subsequently the residual vectors $\mathbf{r}_i$ can be calculated as

$$\mathbf{r}_0 = \begin{bmatrix} y_1 - \hat{y}_1 \\ \vdots \\ y_n - \hat{y}_n \end{bmatrix} \text{ and } \mathbf{r}_j = \begin{bmatrix} \frac{\partial y_1}{\partial x_j} - \frac{\partial \hat{y}_1}{\partial x_j} \\ \vdots \\ \frac{\partial y_n}{\partial x_j} - \frac{\partial \hat{y}_n}{\partial x_j} \end{bmatrix} \text{ for } j = 1..n_{dv} \tag{6.34}$$

A biased estimate of the covariance matrix, which is the best estimate obtainable under the present assumptions, can be calculated (*Khuri and Cornell,* 1996 and *Zellner*, 1962)

$$\hat{\mathbf{C}} = \frac{[\mathbf{r}_0 : \mathbf{r}_1 : \ldots : \mathbf{r}_{n_{dv}}]^T [\mathbf{r}_0 : \mathbf{r}_1 : \ldots : \mathbf{r}_{n_{dv}}]}{n} \tag{6.35}$$

Using $\mathbf{W} = \mathbf{I}_n \otimes \hat{\mathbf{C}}^{-1}$ as in Eq. (6.32), a new set of coefficients can be found which is subsequently used to estimate a new covariance matrix $\hat{\mathbf{C}}$. The procedure is then iterated until convergence is achieved for the coefficients. The following algorithm called the iterative covariance algorithm, is used to solve the iteratively weighted least squares problem and is defined as follows:

i. Assume a form for the error-correlation $\hat{\mathbf{C}}$

ii. Solve the least square problem with initial weight matrix $\mathbf{W} = \mathbf{I}$

   a) Calculate residual vectors $\mathbf{r}_j$, j = 1,..,*n* according to Eq (6.34)

   b) Estimate $\hat{\mathbf{C}}$ from Eq. (6.35)

   c) Calculate |**b**| and if the variation of |**b**| is less than a specified small number than stop, otherwise return to a)

iii. Calculate the t_{stat} and other statistical parameters

Different forms of the estimated covariance submatrix $\hat{\mathbf{C}}$

If the approximating function $\hat{y}$ of Eq. (6.4) used is a polynomial of order P, then the total number of parameters to be estimated is $n_b + n_c$ where

$$n_b = \prod_{i=1}^{P} \frac{(n_{dv} + i)}{i!} \tag{6.36}$$

is the number of coefficients in the polynomial approximating functions and

$$n_c = \sum_{i=1}^{n+1} i = \frac{(n_{dv} + 1)(n_{dv} + 2)}{2} \tag{6.37}$$

is the number of unknown entries in the estimated covariance submatrix $\hat{\mathbf{C}}$. For example, for a cubic polynomial in two design variables there are a total of 10 + 6 =16 terms to be estimated.

In general the component $\hat{\mathbf{C}}$ of the covariance matrix as in Eq. (6.31), may be written as

$$\hat{\mathbf{C}} = \begin{bmatrix} C_{00} & C_{01} & \cdots & C_{0n_{dv}} \\ C_{00} & C_{11} & \vdots & C_{1n_{dv}} \\ \vdots & \vdots & \ddots & \vdots \\ C_{00} & C_{00} & \cdots & C_{n_{dv}n_{dv}} \end{bmatrix} \text{ where } C_{ij} = C_{ji} \tag{6.38}$$

where subscript '0' indicates function values while subscripts [1,..,n_{dv}] denote derivatives. For example C_{05} is the correlation between function values and the derivatives of the 5th design variable.

Based on the physics of the problem or on experience, an a priori knowledge of the structure of $\hat{\mathbf{C}}$ might make it possible to reduce the number of unknowns, while at the same time achieving better accuracy. One simpler form corresponds to all derivatives having the same correlation

$$\hat{\mathbf{C}} = \begin{bmatrix} C_{00} & C_1 & \cdots & C_1 \\ C_1 & C_2 & \vdots & C_3 \\ \vdots & \vdots & \ddots & \vdots \\ C_1 & C_3 & \cdots & C_2 \end{bmatrix} \tag{6.39}$$

An even simpler form arises when derivatives and functions are not correlated, but have different variances. In this case the resulting covariance submatrix $\hat{\mathbf{C}}$ is a diagonal matrix

$$\hat{\mathbf{C}} = \begin{bmatrix} C_{00} & 0 & \cdots & 0 \\ 0 & C_{11} & \vdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & C_{11} \end{bmatrix} \tag{6.40}$$

A special case of the covariance matrix is found when all gradients have the same variance

$$\hat{\mathbf{C}} = \begin{bmatrix} C_{00} & 0 & \cdots & 0 \\ 0 & C_{11} & \vdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & C_{n_{dv},n_{dv}} \end{bmatrix} \tag{6.41}$$

Finally, the greatest simplification possible is $\hat{\mathbf{C}} = \sigma^2 \mathbf{I}_{n_{dv}+1}$, that is the variance of the errors in the function and in the gradients are the same, and no correlation exists between them.

Numerical tests

Numerical examples of approximations obtained from function and gradient data are presented and discussed. The response surface approximations obtained by using function and gradient information were compared with approximations obtained by using function values only. The following numerical examples attempt to find evidence of the

advantages and disadvantages of including gradient information in the fitting procedure used to calculate the parameters of a response surface approximation.

Modeling or Bias Errors

This section is focused on modeling (bias) error, that is when the approximating function has a different form then the true function. It is assumed that the number of design points where function and gradient information are available is fixed, as may happen when availability of computational resources is limited. Consider the case of when the number of function values, n, is sufficient to obtain a polynomial of order P. If gradients are available, the total amount of data available increases from n to n (n_{dv}+1).

The benefits of using gradients to improve the quality of an approximation is demonstrated by the following two example problems where a quadratic response surface approximation is used. In the first example the true function and the approximating function are function of two design variables while in the second example the true functions and the approximations are function of ten design variables. The two examples presented consist of functions of variables that range from -1 to +1.

The form of the response surface approximation sought in the first example is a quadratic polynomial in the two variables x_1 and x_2

$$\hat{y} = (x_1, x_2) = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_1^2 + b_4 x_1 x_2 + b_5 x_2^2 \qquad (6.42)$$

Approximations based only on function values obtained from a 3-level factorial design, for a total of nine points, are compared to the approximations obtained with gradient and function values at the same design points for a total of 27 data.

Example 1. Quadratic approximation of a 4th order polynomial

Consider the case where the true function is the following full quartic polynomial in two variables:

$$\begin{aligned} f(x_1,x_2) = 1 + x_1 + x_2 + x_2^2 + x_1x_2 + x_2^2 + \\ + x_1^3 + x_2x_1^2 + x_1x_2^2 + x_2^2 + \\ + x_1^4 + x_2x_1^3 + x_2^2x_1^2 + x_2^3x_2 + x_2^4 \end{aligned} \tag{6.43}$$

and two quadratic polynomials approximations are created:

- $\hat{y}_g$ approximation obtained using function and gradients at nine design points
- $\hat{y}_f$ approximation obtained using only function values at the same nine design points used to obtain $\hat{y}_g$.

The covariance and weight submatrices for $\hat{y}_g$ after six iterations of the iterative covariance algorithm are:

$$\hat{\mathbf{C}} = \begin{bmatrix} 1.749 & 5.132 & 5.132 \\ 5.132 & 16.764 & 14.669 \\ 5.132 & 14.669 & 16.764 \end{bmatrix} \text{ and } \mathbf{W} = \mathbf{I}_9 \otimes \begin{bmatrix} 13.552 & -2.212 & -2.212 \\ -2.212 & 0.616 & 0.138 \\ -2.212 & 0.138 & 0.616 \end{bmatrix} \tag{6.44}$$

Note that the resulting approximate correlation matrix shows high correlation between function values and gradients.

Figure 27 shows the two response surfaces $\hat{y}_g$ and $\hat{y}_f$ along with the derivatives with respect to x_1. The derivatives x_2 are not shown since they are identical to x_1

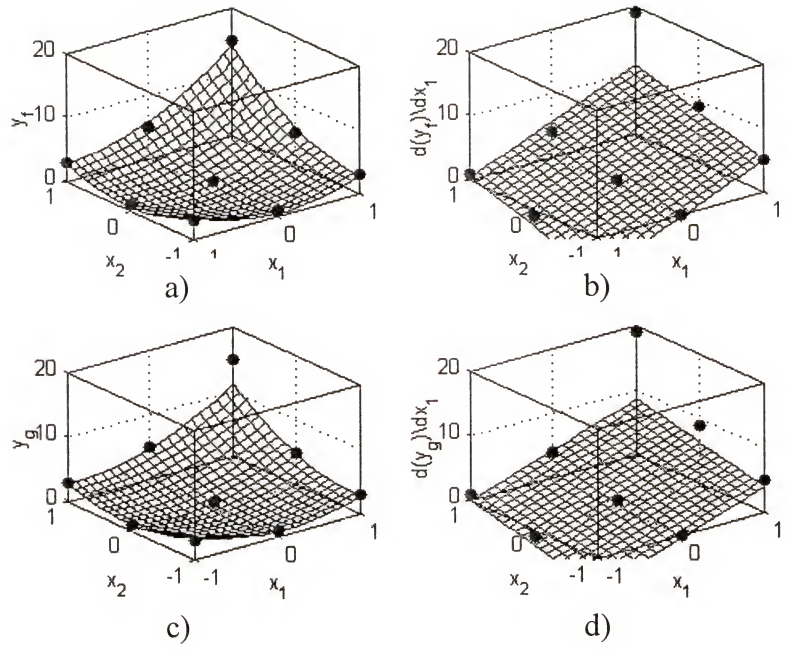


Figure 27: Response surfaces approximations and corresponding derivatives. Dots represent the data used a) response surface obtained from function values only $\hat{y}_f$ b) derivative of $\hat{y}_f$ c) response surface obtained from function values and gradient data $\hat{y}_g$ d) derivative of $\hat{y}_g$

Table 20 presents the value of the coefficients obtained for $\hat{y}_g$ and $\hat{y}_f$ along with the values of the *RMSE* based on function values and gradient values. It should be noted that the *RMSE* presented in Table 20 is calculated from

$$RMSE = \sqrt{\frac{SSE}{n}} \tag{6.45}$$

where *SSE* can be the sum of the square errors on the function or on the gradients depending on the case examined. From the *RMSE* values reported in Table 20, it appears that the approximation based only on function values is more accurate. However when the *RMSE* was checked from the analytical expression, called *Re*, $\hat{y}_g$ resulted in a more accurate approximation than $\hat{y}_f$ in both function values and gradient values. In addition

Table 20 presents the approximation $\hat{y}_{MinBias}$ obtained by minimizing the RMSE over the entire design domain from it analytical formulation and the approximation $\hat{y}_I$ obtained using function and gradient data but using the Identity matrix as a weighting matrix.

Table 20: Parameters and error measures of the approximation $\hat{y}_{MinBias}$ that minimizes analytically the bias error, of the approximation $\hat{y}_g$, obtained using function and gradient values with iterative covariance algorithm, of the approximation $\hat{y}_I$ obtained using function and gradient values with an identity matrix as a weighting matrix and of the approximation $\hat{y}_f$ obtained using only function values. The underscript *f* refers to function values while *g* refers to gradient values. RMSE refers to the original nine points while *Re* refers to the analytical RMSE calculated over the entire design domain

	b_0	b_1	b_2	b_3	b_4	b_5	$RMSE_f$	$RMSE_g$	Re_f	Re_g
$\hat{y}_{MinBias}$	0.717	1.933	1.933	2.190	2.200	2.190	1.294	4.111	0.374	1.574
$\hat{y}_f$	0.556	2.667	2.667	2.667	3.000	2.667	0.588	3.203	0.796	1.515
$\hat{y}_g$	0.623	1.913	1.913	2.187	2.300	2.187	1.323	4.094	0.388	1.573
$\hat{y}_I$	-0.675	3.266	3.266	3.589	3.750	3.589	1.206	2.479	1.466	2.214

The largest gains in accuracy were noted in the function approximation, while the gradient approximation shown in Table 20 did not gain from using gradient data; if anything it got slightly worse. This unexpected result might be explained by Figure 27 b) and d). The true gradients are cubic polynomials of the two variables, while the gradients of the two approximations are only linear. When significant curvature is present, a linear function cannot give good approximation over large intervals. In addition using identity matrix as a weighting matrix gave the most inaccurate results. Moreover, the approximation obtained from the iterative covariance algorithm is almost identical to $\hat{y}_{MinBias}$, the quadratic polynomial that has the minimum analytical RMSE over the entire design domain. This last approximation is obtained by finding the parameters b_i of a

quadratic polynomial of Eq. (6.42) that minimize the analytical expression of the RMSE, *Re* over the entire domain

$$\mathrm{Re} = \frac{\int_V (TrueFunction - \hat{y}_{MinBias})^2 dV}{\int_V dV} \tag{6.46}$$

Increasing the order of the approximation

Next, the effect of using derivative data to increase the order of the approximation was considered. If the number of data points is enough to obtain a quadratic polynomial approximation from function values only, then the number of points n satisfies

$$n > n_{b,P=2} > \frac{(n_{dv}+1)(n_{dv}+2)}{2} \tag{6.47}$$

where $n_{b,P=2}$ is the number of parameters in a quadratic polynomial and n_{dv} is the number of variables of the quadratic polynomial. At each point there are n_{dv} derivatives, so that the total number of data available from n points is

$$n_t > \frac{(n_{dv}+1)(n_{dv}+2)}{2}(n_{dv}+1) \tag{6.48}$$

The number coefficients of a cubic polynomial can be obtained from Eq. (6.48) as

$$n_{b,P=3} = \frac{(n_{dv}+1)(n_{dv}+2)n_{dv}+3)}{6} \tag{6.49}$$

The number of coefficients n_c to be estimated in the covariance matrix $\hat{\mathbf{C}}$ of Eq. (6.38) by the iterative procedure are

$$n_c = \frac{(n_{dv}+1)(n_{dv}+2)}{2} \tag{6.50}$$

It follows that for $n_{dv} > 1$ (i.e., for problems with more than one design variable) there is enough information to fit a cubic polynomial using function and gradient

information when there are enough points to fit a quadratic function using only function values, since

$$n_t > \frac{(n_{dv}+1)(n_{dv}+2)}{2}(n_{dv}+1) > \frac{(n_{dv}+1)(n_{dv}+2)(n_{dv}+3)}{6} + \frac{(n_{dv}+1)(n_{dv}+2)}{2} \quad (6.51)$$

The following examples present the results obtained by creating cubic polynomial approximations using function and gradient information at nine points, that constitute a factorial design in two design variables in the range ± 1. The nine points, the true function as well as the quadratic approximation based on function values only, are the same as in Example 1.

Example 2: Cubic approximation of 4th order polynomial

The case was considered in which the true function is a full quartic polynomial, as in Eq. (6.43), and the two approximations being constructed are:

- $\hat{y}_g$ cubic approximation obtained using the function values and the gradients at nine points.
- $\hat{y}_f$ quadratic approximation obtained using only function values at the same nine design points used for $\hat{y}_g$.

The covariance and weight matrix obtained after seven iterations of the iterative covariance algorithm with a convergence criteria equal to 10^{-4} were

$$\hat{\mathbf{C}} = \begin{bmatrix} 0.1233 & 0 & 0 \\ 0 & 3.6674 & 3.5098 \\ 0 & 3.5098 & 3.6674 \end{bmatrix} \quad \mathbf{W} = \mathbf{I}_9 \otimes \begin{bmatrix} 8.1117 & 0 & 0 \\ 0 & 3.2413 & -3.1019 \\ 0 & -3.1019 & 3.2413 \end{bmatrix} \quad (6.52)$$

It is interesting to note that the correlation matrix shows strong correlation between the gradients, and none between the gradients and the function values, even

though no simplifications were imposed on the form of the correlation matrix. The quadratic approximation obtained using only function values $\hat{y}_f$, is obviously the same one presented earlier in Table 20, and it is reported again for completeness in Table 21.

Table 21: Parameters and error measures of the cubic approximation $\hat{y}_g$, obtained from function and gradient values and the quadratic approximation $\hat{y}_f$ obtained using only function values. The underscript *f* refers to function values and *g* refers to gradient values. RMSE refers to the original nine points and *Re* to the analytical *RMSE*

	B_0	b_1	b_2	b_3	b_4	b_5	b_6	b_7	b_8	b_8
$\hat{y}_f$	0.556	2.667	2.667	2.667	3.000	2.667	-	-	-	-
	$RMSEe_f$	$RMSE_g$	Re_f	Re_g						
	0.588	3.203	0.793	1.685						
	b_0	b_1	b_2	b_3	b_4	b_5	b_6	b_7	b_8	b_9
$\hat{y}_g$	0.023	1.00	1.00	3.066	2.918	3.066	1.0	1.0	1.0	1.0
	$RMSEe_f$	$RMSE_g$	Re_f	Re_g						
	0.351	1.913	0.488	1.089						

Table 21 shows that the cubic approximation based on gradient information is more accurate than the approximation derived from function values only. Moreover comparing the results shown in Table 21 with the results shown in Table 20 it appears that using the gradient information to increase the order of the approximation increases the error in the function predictions, but it reduced by 50% the errors in the predictions of the gradients.

Example 3: Ten design variable case

An additional numerical experiment was carried out using ten design variables. The true function was a full quartic polynomial in the ten design variables with all

coefficients equal to 1.0. Five different approximations were constructed based on the values of 70 design points. The points were chosen using by LHS method:

- $\hat{y}_{g,Quad}$ quadratic approximation obtained using function values and gradients at 70 design points.
- $\hat{y}_{g,Cubic}$ cubic approximation obtained using function values and gradients at 70 design points.
- $\hat{y}_f$ quadratic approximation obtained using only function values at the same 70 design points.
- $\hat{y}_{I,Quad}$ quadratic approximation obtained using function values and gradients at 70 design points and the Identity matrix as a weighting matrix.
- $\hat{y}_{I,Cubic}$ cubic approximation obtained using function values and gradients at 70 design points and the Identity matrix as a weighting matrix.

Note that a quadratic polynomial in 10 design variables has 66 coefficients and therefore the approximation based on function values only, $\hat{y}_f$ will have just about enough information to obtain the 66 parameters. The results obtained are shown in Table 22, where due to their large number, the coefficients of each approximation were omitted. In Table 22 the quadratic and cubic polynomials having the minimum possible bias error are also presented. Note that these two last approximations are guaranteed to have the lowest analytical Re_f for the their class of polynomials, however they are not guaranteed to have the lowest sum of the analytical *RMSE* of the gradients, $Re_{g,Sum}$. As can be seen, the iterative covariance algorithm gives a much more accurate approximations than the approximation based on function values only. However in contrast with the two design

variables case, the approximation obtained using the identity matrix as a weighting matrix gives the best approximation. In order to establish if the identity matrix consistently gives better results for this higher dimensional problem, the same problem was repeated using 100 and 150 design points. For these two last cases the iterative covariance algorithm gave slightly better results than the approximations based on the identity matrix as a weighting matrix, suggesting that as the number of data available increases the iterative covariance matrix algorithm becomes more and more accurate.

Table 22: Response surfaces obtained from the true function and true gradient values at 70 design points. The true function is a quartic polynomial in ten design variables. *Re* indicates the analytical *RMSE* calculated over the whole design domain. The underscript *f* refers to function values and *g* refers to gradient values. The $Re_{g,Sum}$ is the sum of the ten *Re* of the gradients. Min Bias refers to quadratic and cubic polynomials that minimize the analytical bias error of the function.

	Re_f	$Re_{g,Sum}$
$\hat{y}_{g,Quad}$	5.806	55.58
$\hat{y}_{g,Cub}$	4.246	46.18
$\hat{y}_f$	21.62	---------
$\hat{y}_{I,Quad}$	4.532	46.21
$\hat{y}_{I,Cubic}$	4.152	45.80
Min Bias Quad	3.819	50.01
Min Bias Cubic	2.692	39.05

Noise Errors

In this section the primary focus of the investigation is on noise error. Normally distributed noise is added to the values of the true function and then an approximation is fitted using the noisy data. The true function is set to be a quadratic polynomial as

$$f(x_1, x_2) = 1 + x_1 + x_2 + x_2^2 + x_1 x_2 + x_2^2 \qquad (6.53)$$

Random errors from a normal distribution are added to the value of the true function to generate noisy data that are subsequently used to obtain the response surface expressions. Random error is added in the following fashion

$$y(\mathbf{x}_i) = f(\mathbf{x}_i) + \varepsilon_i\,,\ x_i = [x_{1i}, x_{2i}] \text{ and } \varepsilon_i = N(0,\sigma^2) \tag{6.54}$$

$$\frac{\partial y(\mathbf{x}_i)}{\partial x_j} = \frac{\partial f(\mathbf{x}_i)}{\partial x_j} + \varepsilon_i\,,\ x_i = [x_{1i}, x_{2i}] \text{ for } j = 1..2 \text{ and } \varepsilon_i = N(0,\sigma^2) \tag{6.55}$$

The two approximations generated from the data are:

- $\hat{y}_g$ approximation obtained using the function and gradients values at nine design points. The covariance matrix used in the iterative covariance algorithm is the fully correlated covariance matrix of Eq. (6.38).
- $\hat{y}_f$ approximation obtained using only function values at the same nine design points used for $\hat{y}_g$.

Due to the random nature of the error, in order to make a reasonable comparison between the two approximations, different sets of different randomly generated errors were used. The same sets were used to obtain $\hat{y}_g$ and $\hat{y}_f$.

Example 1: Approximation of 2nd order polynomial with random errors $\sigma^2 = 0.03$

In this first example noise with variance $\sigma^2 = 0.03$ is added to the true function and to its derivatives. Ten sets of random errors were generated and added to the values of the true function. $\hat{y}_f$. The results obtained are presented in Table 23 and Table 24. In these tables, the results indicated by Re_{f101} refer to the root error mean square of the approximating function calculated at 101^2 points while Re_{g101} refers to the root mean square error of the gradients at the same 101^2 points.

Table 23: Response surfaces obtained from the true function and true gradient values with added noise. The noise added had variance $\sigma^2 = 0.03$ for both function and gradient values. Errors measures *Re* are calculated at 101^2 check points. The underscript *f* refers to function values and *g* refers to gradient values. The number 1 or 2 refers to the design variable number. Av means average.

Set	b_0	b_1	b_2	b_3	b_4	b_5	$Re_{f,101}$	$Re_{g1,101}$	$Re_{g2,101}$
1	1.0651	1.0135	1.0055	0.9927	0.9863	0.9860	0.0588	0.0179	0.0190
2	0.9193	0.9752	0.9470	1.0515	1.0031	1.0030	0.0726	0.0650	0.0531
3	1.0033	0.9923	0.9867	1.0095	1.0227	0.9592	0.0189	0.0189	0.0511
4	0.9966	1.0350	1.0270	0.9929	1.0375	1.0619	0.0376	0.0421	0.0801
5	1.0033	0.9923	0.9867	1.0095	1.0227	0.9592	0.0189	0.0189	0.0511
6	1.0587	1.0928	1.0728	1.0292	1.0593	1.0049	0.1008	0.1047	0.0808
7	0.9739	1.1044	1.0681	0.9938	1.0205	1.0088	0.0773	0.1053	0.0699
8	1.0283	0.9628	0.9855	1.0246	0.9777	0.9838	0.0407	0.0487	0.0271
9	0.9003	0.9709	1.0318	0.9921	1.0081	1.0481	0.0908	0.0308	0.0647
10	1.0555	0.9141	0.9974	1.0657	0.9957	0.9440	0.0816	0.1151	0.0654
Av.							0.0598	0.0567	0.0562

Table 24 Response surfaces obtained from the true function values with added noise. The noise had variance $\sigma^2 = 0.03$. The errors measures *Re* are calculated at 101^2 check points. The underscript *f* refers to function values and *g* refers to gradient values. The number 1 or 2 refers to the design variable number. Av means average.

Set	b_0	b_1	b_2	b_3	b_4	b_5	$Re_{f,101}$	$Re_{g1,101}$	$Re_{g2,101}$
1	0.9304	1.0155	1.0229	1.1265	1.0548	1.0467	0.0490	0.1517	0.0671
2	1.1521	0.9986	0.9945	0.7644	1.0428	0.9449	0.0921	0.2759	0.0692
3	1.0458	0.9323	1.0761	0.9278	0.8780	0.9396	0.0779	0.1293	0.1258
4	1.1633	1.0106	1.0230	0.9160	0.8722	0.9353	0.1260	0.1236	0.1085
5	1.0458	0.9323	1.0761	0.9278	0.8780	0.9396	0.0779	0.1293	0.1258
6	0.9677	1.0947	0.9920	1.1646	0.8514	1.0044	0.0936	0.2309	0.0872
7	0.8701	0.9703	1.0123	0.9767	0.9470	1.2008	0.0964	0.0508	0.2366
8	1.2869	1.1534	0.9546	0.8953	1.0263	0.7529	0.2083	0.1966	0.2922
9	0.8999	0.9428	1.0425	1.0323	1.0205	1.0062	0.0972	0.0696	0.0448
10	1.1155	0.9277	1.0783	1.0332	0.9389	0.8933	0.1168	0.0894	0.1513
Av.							0.1035	0.1447	0.1308

Comparing Table 23 and Table 24, it appears that response surface approximations obtained combining function and gradient information are more accurate than the approximation obtained using only function values. In particular, using gradient data in the fitting procedure drastically increases the accuracy of the gradients

approximations. Moreover when using both function and gradient data, despite the amount of data being three times larger than the case of approximation obtained from function values only, the average Re_{f101} for the 10 sets of random errors decreased only by 50%.

The iterative covariance algorithm was used in all the cases of regressions based on function and gradient information. Convergence was assumed when none of the **b** coefficients changed by more than 10^{-4}. The most accurate covariance matrix estimated by the procedure was found for design number 5 as

$$\hat{\mathbf{C}} = 10^{-3}\begin{bmatrix} 31.309 & -11.958 & 4.481 \\ -11.958 & 32.522 & -2.040 \\ 4.481 & -2.040 & 26.360 \end{bmatrix} \tag{6.49}$$

The covariance matrix of this case is estimated very well considering that the noise was generated from

$$\hat{\mathbf{C}} = 10^{-3}\begin{bmatrix} 30.0 & 0.0 & 0.0 \\ 0.0 & 30.0 & 0.0 \\ 0.0 & -0.0 & 30.0 \end{bmatrix} \tag{6.56}$$

The worst estimation of the covariance matrix was found for design number 10 as

$$\hat{\mathbf{C}} = 10^{-3}\begin{bmatrix} 18.743 & -5.125 & 0.429 \\ -5.125 & 11.042 & -3.858 \\ 0.429 & -3.858 & 5.838 \end{bmatrix} \tag{6.57}$$

Note that for cases 5 and 10 the corresponding approximation are also best and worst.

Example 2: Approximation of 2nd order polynomial with random errors $\sigma^2 = 0.06$

A comparison between approximations obtained by using function and gradient data and function data only, was repeated for noise generated from normal distribution of variance $\sigma^2 = 0.06$. The true function with no noise was still given by Eq. (6.46). The results obtained for this new case are presented in Table 25 and Table 26. The benefit of using the gradient values in obtaining the approximation to the true function is once more evident. Major gains in accuracy obtained for the gradients expressions are noticeable.

Table 25: Response surfaces obtained from the true function and gradient values with added noise. The noise had variance $\sigma^2 = 0.06$ for both function and gradient values. The errors measures *Re* are calculated at 101^2 check points. The underscript *f* refers to function values and *g* refers to gradient values. The number 1 or 2 refers to the design variable number. Av means average.

Set	b_0	b_1	b_2	b_3	b_4	b_5	$Re_{f,101}$	$Re_{g1,101}$	$Re_{g2,101}$
1	1.0329	1.0270	1.0477	0.9807	1.0516	0.9888	0.0434	0.0462	0.0579
2	0.9241	0.9949	0.9871	0.9649	0.9409	0.9762	0.0992	0.0537	0.0461
3	1.0582	0.9412	0.9240	0.9895	1.2206	1.0209	0.1124	0.1420	0.1514
4	0.8701	0.9802	1.0031	1.0178	0.9028	1.0077	0.1264	0.0635	0.0574
5	0.8647	0.8756	0.9547	1.0087	0.9505	0.9872	0.1579	0.1281	0.0557
Av.							0.1079	0.0867	0.0737

Table 26: Response surfaces obtained from the true function values with added noise. The noise had variance $\sigma^2 = 0.06$. The error measures *Re* are calculated at 101^2 check points. The underscript *f* refers to function values and *g* refers to gradient values. The number 1 or 2 refers to the design variable number. Av means average.

Set	b_0	b_1	b_2	b_3	b_4	b_5	$Re_{f,101}$	$Re_{g1,101}$	$Re_{g2,101}$
1	0.7464	1.0379	1.0411	1.2617	0.7297	1.1423	0.1767	0.3456	0.2325
2	1.0533	0.8266	0.9747	0.8802	0.9809	0.9108	0.1134	0.2229	0.1076
3	0.7905	1.0504	0.9806	1.3331	1.1442	1.0789	0.1380	0.4007	0.1261
4	0.8462	0.9827	1.0161	0.9221	0.8906	1.1422	0.1464	0.1123	0.1784
5	0.7632	0.8037	0.9442	1.0332	0.9970	1.0780	0.2332	0.2001	0.1067
Av.							0.1615	0.2563	0.1502

Example 3: Approximation of 2nd order polynomial with random errors $\sigma^2 = 0.03$ on the function values and $\sigma^2 = 0.06$ on the gradient values

Next the case of added noise with two different variances, one for the function and one for the gradients data, was considered. The random noise added to the true function values had a variance $\sigma^2 = 0.03$, while the noise added to the true value of the gradients had a variance of $\sigma^2 = 0.06$.

The results obtained from the fit that used function and gradient information are presented in Table 27, while the results obtained for the fit obtained using only function values are shown in Table 28. Comparison of Table 27 and Table 28 suggests that there is still an advantage in including gradient information in constructing an approximation even when the noise in the functions and the gradients is large.

Table 27: Response surfaces obtained using function and gradient having different noise levels. The noise added to function values had variance $\sigma^2 = 0.03$ while the noise added to the gradient values had variance of $\sigma^2 = 0.06$. The error measures *Re* are calculated at 101^2 check points. The underscript *f* refers to function values and *g* refers to gradient values. The number 1 or 2 refers to the design variable number. Av means average

Set	b_0	b_1	b_2	b_3	b_4	b_5	$Re_{f,101}$	$Re_{g1,101}$	$Re_{g2,101}$
1	0.8647	0.8756	0.9547	1.0087	0.9505	0.9872	0.1579	0.1281	0.0557
2	1.2342	1.1079	0.9611	0.7532	0.9996	0.8817	0.1533	0.3074	0.1434
3	0.9443	1.0049	1.0617	0.9754	0.9478	1.0855	0.0598	0.0421	0.1212
4	1.1304	0.9948	0.9044	0.9470	1.0046	0.9137	0.1047	0.0620	0.1389
5	0.8984	1.0365	0.8968	1.0640	1.0391	1.0390	0.0960	0.0862	0.1151
Av.							0.1143	0.1251	0.1148

Table 28: Response surfaces obtained using only function values with added noise having variance $\sigma^2 = 0.03$. The error measures *Re* are calculated at 101^2 check points. The underscript *f* refers to function values and *g* refers to gradient values. The number 1 or 2 refers to the design variable number. Av means average.

Set	b_0	b_1	b_2	b_3	b_4	b_5	$Re_{f,101}$	$Re_{g1,101}$	$Re_{g2,101}$
1	0.7632	0.8037	0.9442	1.0332	0.9970	1.0780	0.2332	0.2001	0.1067
2	1.3312	0.9428	0.9263	0.7853	1.0684	0.7551	0.2096	0.2599	0.2976
3	1.0562	1.0651	1.0807	0.8518	0.9753	1.0221	0.0773	0.1853	0.0859
4	1.0887	1.0504	0.9099	0.9600	1.1519	0.9735	0.1043	0.1121	0.1300
5	0.9815	1.0806	0.9999	0.9457	1.0500	1.0763	0.0586	0.1066	0.0937
Av.							0.1366	0.1728	0.1427

Example 4: Approximation of 2nd order polynomial with random errors $\sigma^2 = 0.3$

The case of added noise with a large variance was also considered. The random noise added to the true function values and the gradients had a variance $\sigma^2 = 0.3$. The results obtained from the fit that used function and gradient information are presented in Table 29, while the results obtained for the fit obtained using only function values are shown in Table 30. Comparison of Table 29 and Table 30 suggests that there is still an advantage in including gradient information in constructing an approximation even when the noise in the gradients is large. However the gains in accuracy obtained by including the gradient information progressively reduce as the noise in the gradient information increases.

Table 29: Response surfaces obtained using function and gradient having the same noise levels. The noise added to function values and the gradient had variance $\sigma^2 = 0.3$. The error measures *Re* are calculated at 101^2 check points. The underscript *f* refers to function values and *g* refers to gradient values. The number 1 or 2 refers to the design variable number. Av means average

Set	b_0	b_1	b_2	b_3	b_4	b_5	$Re_{f,101}$	$Re_{g1,101}$	$Re_{g2,101}$
1	0.6191	0.7402	0.9376	1.0195	0.9389	1.1195	0.3706	0.2619	0.1564
2	0.7782	1.2619	1.2247	1.0345	1.1003	1.1433	0.2641	0.2692	0.2798
3	0.7926	0.9909	0.9405	1.1607	0.8721	0.9561	0.1817	0.2013	0.1068
4	1.1437	1.1736	0.5865	0.9638	1.0744	0.9543	0.2862	0.1822	0.4172
5	1.0111	0.9763	0.9581	1.0311	1.0728	0.8705	0.0563	0.0603	0.1584
Av							0.2317	0.1950	0.2237

Table 30: Response surfaces obtained using only function values with added noise having variance $\sigma^2 = 0.3$. The error measures *Re* are calculated at 101^2 check points. The underscript *f* refers to function values and *g* refers to gradient values. The number 1 or 2 refers to the design variable number. Av means average.

Set	b_0	b_1	b_2	b_3	b_4	b_5	$Re_{f,101}$	$Re_{g1,101}$	$Re_{g2,101}$
1	0.0398	0.9450	1.1996	1.5460	0.7064	1.2874	0.7196	0.6616	0.4260
2	1.1419	1.0876	1.0956	0.7411	0.4707	0.8556	0.2151	0.4405	0.3643
3	0.3195	0.9992	1.0494	1.0920	0.8413	1.7517	0.4604	0.1408	0.8829
4	1.2292	1.0840	0.4836	0.9212	1.0386	0.7425	0.3363	0.1265	0.5978
5	1.1450	0.7860	1.2407	0.7718	0.6142	0.8090	0.2463	0.4090	0.3977
Av							0.3955	0.3556	0.5337

Example 5: Approximation of 2nd order polynomial with correlated random errors

Finally the case of correlated errors was considered. The correlation used to generate the noise to be added to the function and gradient values at each design point was such that the correlation between the function values and the first gradient was equal to 0.3, the correlation between the function and the second gradient was equal to 0.5 and the correlation between the gradients was equal to 0.4. Ten sets of random errors with variances equal to 0.05, 0.005 and 0.1 for the function, first gradient and second gradient, respectively were generated. The results obtained are presented in Table 31, where in this

case *Re* indicates the analytical root mean square error. As can be seen, using the iterative covariance matrix algorithm usually gave the most accurate results. In particular the results obtained from using only function values were always much less accurate than the results obtained using the iterative covariance algorithm. Using the gradient data was useful in terms of accuracy even when no weighting was used, that is the weighting matrix was equal to the identity matrix.

Table 31: Accuracy of the response surfaces obtained from the true function and true gradient values with correlated added noise. Analytical *RMSE* of the approximation based only on function values, $\hat{y}_f$, of the approximation based on function values and gradients using the iterative covariance algorithm $\hat{y}_g$, and of the approximation based on function values and gradients using the identity matrix as a weighting matrix, $\hat{y}_I$.

Set	$\hat{y}_g$	$\hat{y}_f$	$\hat{y}_I$
1	0.105	0.320	0.153
2	0.128	0.196	0.173
3	0.182	0.199	0.181
4	0.045	0.153	0.103
5	0.130	0.204	0.189
6	0.068	0.097	0.066
7	0.021	0.114	0.053
8	0.054	0.125	0.078
9	0.032	0.105	0.070
10	0.026	0.203	0.067
Av.	0.079	0.172	0.113

Reduction of terms in the approximations

The case of a true function that is described by a quadratic polynomial in two design variables, x_1 and x_2, as in Eq. (6.46) was considered and normally distributed random errors as in Eq. (6.47) and Eq. (6.48) are added to the values of the true function at the design points. The approximation created are

- $\hat{y}_g$ cubic approximation obtained using the function values and the gradients at nine design points.
- $\hat{y}_f$ quadratic approximation obtained using only function values at the same nine design points used for $\hat{y}_g$.

The cubic polynomial approximations obtained based on function and gradients values for this case had coefficients very similar to the coefficients shown in Table 23 for the quadratic part and coefficients of the order of 0.005 for cubic terms.

After approximating the 3rd order polynomial on the data set, it was found that the t_{stat} as defined by Eq. (6.30) for third order terms were all in the range 0.05 - 0.7. An elimination procedure based on discarding the term with the lowest t_{stat} was initiated and was continued until all remaining terms all had t_{stat} higher than 2.0. The procedure was applied to all data sets shown in Table 23 and always ended by eliminating all cubic terms. Therefore, the results obtained are equivalent to those one presented in Table 23.

Conclusion on including gradient information in constructing approximations

Using simple mathematical examples, it has been shown that the inclusion of gradient information can be very useful in improving the quality of the approximations. Some cases have been met where using the identity matrix as a weighting matrix produced slightly better results than using the proposed algorithm. However in many more cases the iterative covariance algorithm obtained much more accurate approximations than any other method considered. In particular, the gradients of the response surface approximations built using function and gradient information were more accurate than the gradient of response surfaces that were constructed using only function values. On the other hand based on the much larger amount of data that is available when

gradients are considered the accuracy in the function values obtained from the approximations achieved was less than expected.

Further investigations are required in the field of approximations based on gradient and function information. A generalization of the methods found in analysis of modeling multiple responses as in *Khuri and Cornell* (1996) and in *Robinson* (2000) is needed.

CHAPTER 7
CONCLUDING REMARKS

The use of response surface approximations in expensive structural optimization problems was investigated in this study. The use of response surface approximations in practical problems is that the cost of constructing the approximations is subject to curse of dimensionality. The curse of dimensionality means that the number of data points required to construct a response surface approximation increases fast with the number of design variables. Response surface approximations are thus limited to practical application that have only few design variables.

In this study, methods that extend the applicability of response surface techniques to problems of higher dimensions were explored. Two methods were proposed: correction response surfaces and response surfaces based on gradient values.

Correction response surface techniques were applied successfully to two panel structural design examples. One example was the optimization of composite hat-stiffener panel subjected to stress and buckling constraints for a blended-wing-body transport. Coupling a single term correction response surface with low-fidelity methods in the form of closed form solutions reduced the computational effort needed to obtain an optimal design. A second example was the optimization problem of a composite bladed stiffener wing panel for a high speed civil transport. Linear correction response surfaces for the ratio and the difference between low-fidelity and high-fidelity methods of analysis were used. It was shown that for a given amount of computational resources, correction response surfaces were more accurate than traditional response surfaces.

Gradients are often cheaper to calculate than functions in structural optimization problems. The possibility of including gradient information in the construction of response surface approximations to structural responses was investigated. A weighted least square procedure that uses gradients and function information with different levels of noise was obtained. Simple numerical models were used to investigate when the inclusion of gradient information in the construction of a response surface is beneficial.

Including gradient information in the construction of the response surface approximation resulted in large gains in the accuracy of the function approximation when only modeling errors were considered. Even when only noise errors were considered, response surface approximations obtained combining function and gradient information resulted to be more accurate than the approximation obtained using function values only. The increased accuracy obtained by using gradient information for the response surface approximations in the presence of noise errors diminished as the magnitude of the noise increased. Moreover by using function and gradients values it was possible to increase the order of the polynomial approximations obtainable from the data. The higher order polynomial approximations were more accurate than the ones obtained by using lower order polynomial approximations based on function values only.

LIST OF REFERENCES

Arora, J. S. Introduction to optimum design. New York, McGraw-Hill, 1989.

Balabanov, V., Giunta, A. A., Golovidov, O., Grossman, B., Mason, W. H., Watson, L. T. and Haftka, R. T. "A reasonable design space approach to response surface approximations." *Journal of Aircraft*, Vol. 36, No. 1 January-February, 1999, pp. 308-315.

Balabanov, V., Haftka, R. T., Grossman, B., Mason, W. H. and Watson, L. T. "Multifidelity response surface model for HSCT wing bending material weight". 7th AIAA/USAF/NASA/ISSMO Symposium on multi disciplinary analysis and optimization, St. Louis MO, 1998, pp. 778-788.

Balabanov, V., Kaufman, M., Giunta, A. A., Haftka, R. T., Grossman, B., Mason, W. H. and L.T. Watson. "Developing customized wing weight function by structural optimization on parallel computers". 37th AIAA/ASME/ASCE/AHS/ASC Structures Dynamics, and Material Conference, Salt Lake City, UT, 1996, Vol. 1, pp. 113-125.

Barthelemy, J. F. M. and Haftka, R. T. "Approximation concepts for optimum structural design-a review." *Structural Optimization*, Vol. 5, 1993, pp. 129-144.

Box, G. E. P. and Behnken, D. W. "Some new three level designs for the study of quantitative variables." *Technometrics*, Vol. 2, 1960, pp. 195-241.

Box, G. E. P. and Draper, N. R. "A basis for the selection of a response surface design." *Journal of American Statistical Association*, Vol. 54, 1959, pp. 622-654.

Box, G. E. P. and Draper, N. R. Empirical model building and response surface. New York, John Wiley & Sons, 1987.

Box, G. E. P. and Hunter, J. S. "Multifactor experimental design for exploring response surface." *Annals of Mathematical Statistics.*, Vol. 28, 1957, pp. 195-241.

Box, G. E. P. and Wilson, K. B. "On the experimental attainment of optimum conditions (with discussion)." *Journal of Royal Statistical Society*, Vol. B13, 1951, pp. 1-45.

Box, M. J., Hunter, W. G., Mac Greagor, J. F. and Erjavec, J. "Some problems associated with the analysis of multiresponse data." *Technometrics*, Vol. 15, 1970, pp. 33-51.

Burgee, S., Giunta, A. A., Balabanov, V., Grossman, B., Mason, W. H., Narducci, R., Haftka, R. T. and Watson, L. T. "A coarse-grained parallel variable-complexity multidiscipliary optimization paradigm." *The International Journal of supercomputing applications and high performance computing*, Vol. 10, No. 4, 1996, pp. 269-299.

Bushnell, D., PANDA2, "Interactive program for minimum weight design of stiffened cylindrical panels and shells" *Computers and Structures*, Vol. 16, 1983, pp. 167-185.

Campen, D. H. V., Nagtegaal, R. and Schoofs, A. J. G. Approximation methods in structural optimization using experimental design for multiple responses. "Multicriteria design optimization". J. K. H. Eschenauer, A. Oscyczka. Berlin, Germany, Springer-Verlag, 1995, pp. 205-228.

Chang, K. J., Haftka, R. T., Giles, G. L. and Kao, P. J. "Sensitivity-based scaling for approximating structural response." *Journal of Aircraft*, Vol. 30, No. 2, 1993, pp. 283-287.

Dantzig, G. Linear programming and extensions. Princeton, New Jersey, Princeton University Press, 1963.

Dyck, D. N. and Lowther, D. A. "Response surface modeling of magnetic device performance using function value and gradient." *International Journal of Applied Electromagnetism and Mechanics*, Vol. 9, 1998, pp. 241-248.

Eldred, M. S. and Schimel, B. D. "Extended parallelism models for optimization on massively parallel computers". 3rd World Congress of Structural and Multidisciplinary Optimization, Buffalo, New York, 1999, pp. 215-217.

Etman, L. F. P. (1992). "Some global and mid-range approximation concepts for optimum structural design". Eindhoven, Eindhoven University of Technology.

Etman, L. F. P., Adriaenes, J. M. T. A., Slagmaat, M. T. P. V. and Schoofs, A. J. G. "Crash worthiness design optimization using multipoint sequential linear programming." *Structural Optimization*, Vol. 12, 1996, pp. 222-228.

Etman, L. F. P., Campen, D. H. V. and Schoofs, A. J. G. "Optimization of multibody system using approximation concepts". IUTAM Symposium on Optimization of Mechanical Syastems, Dordrecht, 1996, pp. 81-88.

Fadel, G. M., Ridley, M. F. and Barthelemy, J. F. M. "Two-point exponential approximation method for structural optimization." *Structural Optimization*, Vol. 2, No. 2, 1990, pp. 117-124.

Fedrov, V. V. Theory of optimal experiments. New York, Academic Press, 1972.

Finney, D. J. "The fractional replication of factorial arrangements." *Annals of Eugenics*, Vol. 12, 1945, pp. 291-301.

Fleury, C. "Efficient approximation concepts using second order information." *International Journal of Numerical Methods in Engineering*, Vol. 28, 1989, pp. 2041-2058.

Free, J. W., Parkinson, A. R., Bryce, G. R. and Balling, R. J. "Approximations of computationally expensive and noisy functions for constrained nonlinear optimization." *Journal of Mechanism, Transmission and Automation in Desing*, Vol. 109, 1987, pp. 528-523.

Giunta, A. A., Balabanov, V., Burgee, S., Grossamn, B., Haftka, R. T. and Mason, W. H. Variable-complexity multi-disciplinary design optimization using parallel computers. Berlin, Germany, Springer, 1995.

Giunta, A. A., Dudley, J. M., Narducci, R., Grossman, B., Haftka, R. T., Mason, W. H. and Watson, L. T. "Noisy aerodynamic response and smooth approximations in HSCT design". 5th AIAA/USAF/NSAS/ISSMO Symposium on Multidisciplinary Analysis and Optimization, Panama City, Florida, 1994, pp. 1117-1128.

Goel, A., Baker, C. A., Shaffer, C. A., Grossman, B., Mason, W. H., L.T.Watson and Haftka, R. T. (1999). "Vizcraft: a problem solving environment for configuration design of a high speed civil transport". Blacksbourg, Virginia, Virginia Polytechnic Institute and State University.

Haftka, R. T. "Combining global and local approximations." *AIAA Journal*, Vol. 29, No. 9, 1991, pp. 1523-1525.

Haftka, R. T., Nachlas, J. A., Watson, L. T. and Desai, R. "Two-point constraint approximation in structural optimization." *Computer Methods in Applied Mechanics and Engineering*, Vol. 60, No. 3, 1987, pp. 289-301.

Haftka, R. T. and Shore, C. P. (1979). "Approximation method for combined thermal/structural design", NASA.

Haug, E. J. and Aurora, J. S. Applied optimal design. New York, John Wiley & Sons, 1979.

Hill, W. J. and Hunter, W. G. "A review of response surface methodology: a literature survey." *Technometrics*, Vol. 8, No. 4, 1966, pp. 571-590.

Ho, Y. C., Shi, L., Dai, L. and Gong, W. "Optimizing discrete event dynamic systems via the gradient surface method". 30th IEEE Conference on decision and control, Piscataway, New Jersey, 1991, Vol. 1, pp. 104-109.

Houten, M. H. V., Schoofs, A. J. G. and Campen, D. H. V. "Response surface techniques in structural optimization". 1st WCSMO World Congress of Structural and Multidisciplinary Optimization, Germany, 1995, pp. 89-95.

Kaufman, K., Balabanov, V., Burgee, S. L., Giunta, A. A., Grossman, B., Mason, W. H. and Watson, L. T. "Variable complexity response surface approximation for wing structural weight in HSCT design." *Computational Mechanics*, Vol. 18, 1996, pp. 112-126.

Keulen, F. V., Liu, B. and Haftka, R. T. "Noise and discontinuity issues in response surfaces based on functions and derivatives". AIAA-2000-1363, 41th AIAA/ASME/ASCE/AHS/ASC/ Structures, Structural Dynamics and Material Conference 2000,,

Keulen, F. V., Polynkine, A. A. and Toropov, V. V. "Shape optimization with adaptive mesh refinement: target error selection strategies." *Engineering Optimization*, Vol. 28, 1997, pp. 95-125.

Keulen, F. V., Toropov, V. V. and Polynkine, A. A. "Shape optimization strategies using the multi-point approximation method and adaptive mesh refinement". 1st World Congress of Structural and Multidisciplinary Optimization, 1995, pp. 67-74.

Khuri, A. I. and Cornell, J. A. Response Surfaces, 2nd Edition. New York, Marcel Dekker, INC., 1996.

Kiefer, J. "Optimal experimental design." *Journal of the Royal Statistical Society, Series B*, Vol. 21, 1959, pp. 272-304.

Kiefer, J. "Optimum designs in regression problems." *Annals of Mathematical Statistics*, Vol. 32, 1961, pp. 298-325.

Kiefer, J. and Wolfowitz, J. "Optimum designs in regression problems." *Annals of Mathematical Statistics*, Vol. 30, No. 271-294, 1959.

Knill, D. L., .Giunta, A. A., Baker, C. A., Grossman, B., Mason, W. H., Haftka, R. T. and Watson, L. T. "Response surface model combining linear and euler aerodynamics for super sonic transport design." *Journal of Aircraft*, Vol. 36, No. 1, 1999, pp. 75-86.

Kok, S., Stander, N. and Roux, W. "Thermal optimization in transient thermoelasticity using response surface approximations." *International Journal for Numerical Methods in Engineering*, Vol. 43, 1998, pp. 1-21.

Krasteva, D. T., Watson, L. T., Baker, C., Grossman, B., Mason, W. H. and Haftka, R. T. "Distributed control parallelism in multidisciplinary design." *Concurrency: Practice and Design*, Vol. 11, No. 8, 1999, pp. 435-459.

Kroo, I., Altus, S., Braun, R., Gage, P. and Sobieski, I. "Multidisciplinary optimization methods for aircraft preliminary design". 5th AIAA/USAF/NASA/ISSMO Symposium on multidisciplinary analysis and optimization, Panama City, Florida, 1994, pp. 697-707.

Lancaster, P. and Salkauskas, K. Curve fitting and surface fitting. London, Academic Press, 1986.

Liu, B. and Haftka, R. T. "Two-level composite wing structural optimization using response surfaces". 7th AIAA/USAF/NASA/ISSMO Symposium on multidisciplinary analysis and optimization, St. Louis, Missouri, 1998

Malkov, V. P. and Toropov, V. V. "Simulation approach to structural optimization using design sensitivities analysis". Engineering optimization in design process, Karlsruhe, Germany, 1991, Vol. 63, pp. 225-231.

Malone, B. and Woyak, S., MODELCENTER Phoenix Integration, Blacksbourg, Virginia

Malone, B. and Woyak, S. "An integration environment to support multidisciplinary design optimization". 3rd World Congress of Structural and Multidisciplinary Optimization, Buffalo, New York, 1999, pp. 289-291.

Mason, B. H., Haftka, R. T., Johnson, E. R. and Farley, G. L. "Variable complexity design of composite fuselage frames by response surface techniques." *Thin Wall Structures*, Vol. 32, No. 4, 1994, pp. 235-261.

McKay, M. D., Conover, W. J. and Beckman, R. J. "A comparison of three methods for selecting values of input variables in the analysis of output from a computer code." *Technometrics*, Vol. 221, 1979, pp. 239-245.

Microsoft Corporation, EXCEL 5.0, 1998

Mitchell, T. J. "An algorithm for the construction of D-Optimal experimental designs." *Technometrics*, Vol. 2, No. 16, 1974, pp. 203-210.

Murthy, D. V. and Haftka, R. T. "Approximation to eigenvalues of modified general matrices." *Composites & Structures*, Vol. 29, 1988, pp. 903-917.

Myers, R. H. and Montgomery, D. C. Response Surface Methodology. New York, Wiley and Sons, 1995.

NASA Contracts NAS1 18862, and 20546.

Papila, M. and Haftka, R. T. "Uncertainty and wing structural weight approximations". 40th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Material Conference, St. Louis, Missouri, 1999, Vol. 2, pp. 998-1002.

Plackett, L. P. and Burman, J. P. "Design of optimum multifactorial experiments." *Biometrica*, Vol. 33, 1946, pp. 305-325.

Ragon, S. A., Gurdal, Z., Haftka, R. T. and Tzong, T. J. "Global/Local structural wing design using response surface techniques". 38th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Material Conference, Kissimmee, Florida, 1997, pp. 1204-1214.

Rankin, C. C., Brogen, F. A., Loden, W. A. and Cabiness, H. D., STAGS 3.0, Lockheed Martin Missiles and Space Co, Inc, Advanced Technology Center, Palo Alto, California

Rasmussen, J. (1996). "Nonlinear programming by cumulative approximation refinement". Aalborg, Denmark, Institute of Mechanical Engineering, Aalborg University.

Riggs, H. R., Tesser, A. and Chu, H. "C^1-continuos stress recovery in finite element analysis." *Computational Methods in Applied Mechanics Engineering*, Vol. 143, 1997, pp. 299-316.

Robinson, K. S. (2000). "Quantile dispersion graphs for design comparison for logistic models and other modeling issues". Doctoral thesis, department of statistics. Gainesville, University of Florida.

Rodriguez, J. F., Renault, J. E. and Watson, L. T. "Convergence using variables fidelity approximation data in a trust region managed augmented lagrangian approximate optimization". 7th AIAA/USAF/NSAS/ISSMO Symposium on multidisciplinary analysis and optimization, St. Loius, Missouri, 1998, pp. 749-768.

Roux, W. J., Stander, N. and Haftka, R. T. "Response surface approximations for structural optimization." *International Journal of Numerical Methods in Engineering*, Vol. 42, No. 1998, pp. 517-534.

SAS Institute, JMP 3.2, Drive Cary, North Carolina, 1998

Schmit, L. A. and Farshi, B. "Optimum design of laminated fiber composite plates." *International Journal of Numerical Methods in Engineering*, Vol. 11, 1977, pp. 623-640.

Schmit, L. A. and Miura, H. (1976). "Approximation concepts for efficient structural synthesis", *AIAA Journal*, Vol 20, 1981, pp 138-147.

Schoofs, A. J. G., Klink, M. B. M. and Campen, D. H. V. "Approximation of structural optimization problems by means of designed numerical experiments." *Structural Optimization*, Vol. 4, 1992, pp. 206-212.

Sellar, R. S., Stelmack, M. A., Batill, S. M. and Renaud, J. E. "Response surface approximations for discipline coordination in multidisciplinary design optimization". 37th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Material Conference, Salt Lake City, Utha, 1996, pp. 583-593.

Sen, A. and Srivastava, M. Regression analysis, theory, methods, and applications. New York, Springer-Verlag, 1990.

Sharifzadeh, S., Koehler, J. R., Owen, A. B. and Shott, J. D. "Using simulators to model transmitted variability in IC manufacturing." *IEEE Transactions on Semiconductor Manufacturing*, Vol. 2, No. 3, 1989, pp. 82-93.

Sobieszczanski-Sobieski, J. and Haftka, R. T. "Multidisciplinary aerospace design and optimization: survey of recent developments." *Structural Optimization*, Vol. 14, No. 1, 1997, pp. 12-33.

Starnes, J. H. and Haftka, R. T. "Preliminary design of composite wings for buckling, stress, and displacement constraints." *Journal of Aircraft*, Vol. 16, 1979, pp. 564-570.

Svanberg, K. "The method of moving asymptotes - a new method for structural optimization." *International Journal for Numerical Methods in Engineering*, Vol. 24, 1987, pp. 359-373.

Tai, J. C., Mavris, D. N. and Schrage, D. P. "Application of a response surface method to the design of tip jet driven stopped rotor/wing concepts". 1st AIAA Aircraft engineering, technology and operations congress, Los Angeles, California, 1995

Tesser, A., Riggs, H. R., Freese, C. E. and Cook, G. M. "Improved variational method for finite element recovery and a posteriori error estimation." *Computational Methods in Applied Mechanical Engineering*, Vol. 155, No. 1998, pp. 15-30.

TNO Road-Vehicles Research Institute, MADYMO 5.1, Delft, Netherlands, 1994

Toomer, C. A., Topliss, M. E. and Hills, D. P. "Aerodynamics optimization using analytic descriptions of the design space." *Journal of Aircraft*, Vol. 35, No. 6, 1998, pp. 882-890.

Toropov, V., Keulen, F. V., Markine, V. and Boer, H. d. "Refinements in the multi-point approximation methods to reduce the effects of noisy structural responses". 6th AIAA/USAF/NASA/ISSMO Symposium on

multidisciplinary analysis and optimization, Bellevue, Washington, 1996, pp. 941-951.

Toropov, V. V. "Simulation approach to structural optimization." *Structural Optimization*, Vol. 1, No. 1, 1989, pp. 37-46.

Toropov, V. V., Filatov, A. A. and Polynkin, A. A. "Multiparameter structural optimization using FEM and multipoint explicit approximations." *Structural Optimzation*, Vol. 6, 1993, pp. 7-14.

Toropov, V. V., Keulen, F. V., Markine, V. L. and Alvarez, L. F. "Multipoint approximation based on response surface fitting: a summary of recent developments". 1st ASMKO UK/ ISSMO, Ilkley, UK, 1999, pp. 371-380.

Toropov, V. V. and Markine, V. L. "The use of simplified numerical models as mid-range approximation". 6th AIAA/USAF/NASA/ISSMO Symposium on multidisciplinary analysis and optimization 1996

Unger, E. R., Hutchinson, M. G., Rais-Rohani, M., Haftka, R. T. and Grossman, B. "Variable-complexity multidisciplinary design of a transport wing." *International Journal of System Automation: Research and Application*, Vol. 2, No. 2, 1992, pp. 87-113.

Vaidya, R. S., Klung, J. C. and Sun, C. T. "Effect of ply thickness and crack tip damage on failure of notched composite laminates". AIAA/ASME/ASCE/ AHS/ASC Structures, Structural Dynamics and Material Conference, Kissimmee, Florida, 1997, Vol. 3, pp. 2224-2237.

Vaidya, R. S. and Sun, C. T. "Fracture criterion for notched composite laminates". AIAA/ASME/ASCE/ AHS/ASC Structures, Structural Dynamics and Material Conference, Salt Lake City, Utah, 1996, Vol. 1, pp. 331-338.

Vanderplaats, G. N. "Efficient algorithm for numerical airfoil optimization." *Journal of Aircraft*, Vol. 16, 1979, pp. 842-847.

Vanderplaats, G. N. "Structural design optimization: status and direction." *Journal of Aircraft*, Vol. 36, No. 1, 1999, pp. 11-20.

Venkataraman, S. and Haftka, R. T. "Integration of finite element analysis program and panel design program". 38th AIAA/ASME/ASCE/AHS/ASC Structures Structural Dynamics and Material Conference, Kissimmee, Florida, 1997, pp. 1215-1224.

Venkataraman, S. and Haftka, R. T. "Design of shell structures for buckling using correction response surface approximations". 7th AIAA/USAF/NASA/ISSMO Symposium on multidisciplinary analysis and optimization, St. Louis, MO, 1998

Venter, G., Haftka, R. T. and J.H. Starnes, J. "Construction of response surface approximations for design optimization". 6th AIAA/NASA/ISSMO Symposium on Multidisciplinary Analysis and Optimization, Bellevue, Washington, 1996, Vol. 1, pp. 548-564.

Vitali, R., Haftka, R. T. and Sankar, B. V. "Correction Response Approximation for Stress Intensity Factors for Composite Stiffened Plates". AIAA/ASME/ASCE/ AHS/ASC Structures, Structural Dynamics and Material Conference, Long Beach, California, 1998, Vol. 4, pp. 2917-2922.

Vitali, R., Haftka, R. T. and Sankar, B. V. "Multifidelity design of stiffened composite panel with a crack". 3rd WCSMO World Congress of Structural and Multidisciplinary Optimization, Buffalo, New York, 1999

Vitali, R., Park, O., Haftka, R. T. and Sankar, B. V. "Structural Optimization of a Hat Stiffened Panel by Response Surface Techniques". 38th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Material Conference, Kissimmee, FL, 1997, Vol. 4, pp. 2983-2993.

VMA Engineering, GENESIS 4.0, Colorado Springs, Colorado, 1997

Wang, L. and Grandhi, R. V. "Design of frame structures using multivariate spline approximation." *AIAA Journal*, Vol. 32, No. 10, 1994, pp. 2090-2098.

Wang, L. and Grandhi, R. V. "Structural reliability optimization using an efficient safety index calculation procedure." *International Journal for Numerical Methods in Engineering*, Vol. 38, 1995, pp. 1721-1738.

Xu, S. and Grandhi, R. G. "Structural optimization with thermal and mechanical constraints." *Journal of Aircraft*, Vol. 36, No. 1, 1999, pp. 29-35.

Zellner, A. "An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias." *Journal of the American Statistical Association*, Vol. 57, No. 298, 1962, pp. 348-368.

BIOGRAPHICAL SKETCH

Roberto Vitali was born in Morbegno in the Alps of Northern Italy in 1967. He graduated from the Civil Engineering department of the University of Pavia (Italy) and worked for a while as Structural Engineer in the architectural firm of his sister in Italy. Tired of building little chalets for skiing vacationers in the Alps, in 1995 he joined the Multidisciplinary and Design Group of Prof. Haftka at the Aerospace Engineering, Mechanics and Engineering Science of the University of Florida. He has been in Gainesville since then and now he really feels like it is time to get out of here.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Raphael T. Haftka, Chair
Distinguished Professor of Aerospace Engineering, Mechanics and Engineering Science Department

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Bhavani V. Sankar, Co-chair
Professor of Aerospace Engineering, Mechanics and Engineering Science Department

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Peter G. Ifju
Professor of Aerospace Engineering, Mechanics and Engineering Science Department

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Loc Vu-Quoc,
Professor of Aerospace Engineering, Mechanics and Engineering Science Department

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Andre I. Khuri
Department of Statistics

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Evgeni Petrov
Visiting Scientist Imperial College
Science, Technology & Medicine of
London

This dissertation was submitted to the Graduate Faculty of the College of Engineering and to the Graduate School and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

December 2000

M. J. Ohanian
Dean, College of Engineering

Winfred M Phillips
Dean, Graduate School